LARGE DEVIATIONS FOR DIFFUSIONS INTERACTING THROUGH THEIR RANKS

AMIR DEMBO*, MYKHAYLO SHKOLNIKOV*, S. R. SRINIVASA VARADHAN†, AND OFER ZEITOUNI‡

Abstract. We prove a Large Deviations Principle (LDP) for systems of diffusions (particles) interacting through their ranks, when the number of particles tends to infinity. We show that the limiting particle density is given by the unique solution of the approriate McKean-Vlasov equation and that the corresponding cumulative distribution function evolves according to the porous medium equation with convection. The large deviations rate function is provided in explicit form. This is the first instance of a LDP for interacting diffusions, where the interaction occurs both through the drift and the diffusion coefficients and where the rate function can be given explicitly. In the course of the proof, we obtain new regularity results for a certain tilted version of the porous medium equation.

1. Introduction

Recently, systems of diffusion processes (particles) interacting through their ranks have received much attention. For a fixed number of particles $N \in \mathbb{N}$, these are given by the unique weak solution of

(1.1)
$$\mathrm{d}X_i(t) = \sum_{j=1}^N b_j\,\mathbf{1}_{\{X_i(t)=X_{(j)}(t)\}}\,\mathrm{d}t + \sum_{j=1}^N \sigma_j\,\mathbf{1}_{\{X_i(t)=X_{(j)}(t)\}}\,\mathrm{d}W_i(t), \quad i = 1,2,\dots,N,$$

where $b_1, b_2, \dots, b_N$ are arbitrary real constants, $\sigma_1, \sigma_2, \dots, \sigma_N$ are arbitrary positive constants, $W_1, W_2, \dots, W_N$ are independent standard Brownian motions and $X_{(1)}(t) \le X_{(2)}(t) \le \dots \le X_{(N)}(t)$ are the ordered particles at time t.

The existence and uniqueness of the weak solution to (1.1) was shown in the work [3], which was motivated by questions in filtering theory, and the system (1.1) has reappeared in the context of stochastic portfolio theory under the name *first-order market model* (see the book [14] and the survey article [15]). Due its central role

Date: November 26, 2012.

2000 *Mathematics Subject Classification.* 60F10, 60H10, 35K55, 82C22.

Key words and phrases. Large deviations, interacting diffusions, porous medium equation, McKean-Vlasov equation, stochastic portfolio theory.

*Research partially supported by NSF grant #DMS-1106627.

†Research partially supported by NSF grant #DMS-1208334.

‡Research partially supported by NSF grant #DMS-0804133 and Israel Science Foundation grant #111/11.

in the analysis of capital distributions in financial markets and long-term portfolio performance therein, as well as its intriguing mathematical features, the ergodicity and sample path properties of the model have undergone a detailed analysis in the case that the number of particles is fixed (see [6], [7], [19], [20] and [21]). Moreover, concentration properties of the solution to (1.1) for large values of N have been studied in [33] and an analogous infinite particle system has been constructed and analyzed in [32].

In the article [37], it was observed that the system of SDEs (1.1) can be rewritten as

$$\mathrm{d}X_i(t) = b(F_{\rho^N(t)}(X_i(t)))\,\mathrm{d}t + \sigma(F_{\rho^N(t)}(X_i(t)))\,\mathrm{d}W_i(t), \quad i = 1,2,\ldots,N, \tag{1.2}$$

where $\rho^N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)}$ is the *empirical measure* of the particle system at time t, $F_{\rho^N(t)}$ is the corresponding cumulative distribution function, and $b:\ [0,1]\to\mathbb{R}$, $\sigma:\ [0,1]\to(0,\infty)$ are functions satisfying $b\big(\frac{j}{N}\big)=b_j$, $\sigma\big(\frac{j}{N}\big)=\sigma_j$ for all $j=1,2,\ldots,N$. The representation (1.2) allows to view the particle system (1.1) as a system of diffusion processes interacting through their mean-field and gives rise to questions on the large N behavior of the empirical measure in (1.2). The significant mathematical challenge here comes from the discontinuity of the diffusion coefficients in (1.2), already at the level of a law of large numbers addressed in [37]. The latter is obtained in [37] under the assumption that the process of spacings between consecutive ordered particles in (1.2) evolves under its stationary distribution (in particular, it was assumed that b is such that the stationary distribution exists). In this case, it was shown that the limiting particle density γ follows the *McKean-Vlasov equation*

$$\int_{\mathbb{R}} f\,\mathrm{d}\gamma(t) - \int_{\mathbb{R}} f\,\mathrm{d}\gamma(0) = \int_0^t \int_{\mathbb{R}} \big[b(F_{\gamma(s)}(\cdot))f' + \frac{1}{2}\sigma(F_{\gamma(s)}(\cdot))^2 f''\big]\,\mathrm{d}\gamma(s)\,\mathrm{d}s \tag{1.3}$$

for all Schwartz functions f and $t\in\mathbb{R}$, and that the corresponding cumulative distribution functions $R=F_{\gamma(\cdot)}(\cdot)$ evolve according to the *porous medium equation with convection* (see the book [42] and the references therein for a thorough treatment)

$$R_t = (\Sigma(R))_{xx} + (\Theta(R))_x, \tag{1.4}$$

where $\Sigma(\cdot)=\int_0^{\cdot}\frac{1}{2}\sigma^2(u)\,\mathrm{d}u$ and $\Theta(\cdot)=\int_0^{\cdot} b(u)\,\mathrm{d}u$.

In this paper, we prove a Large Deviations Principle (LDP) for the sequence $\{\rho^N,\ N\in\mathbb{N}\}$ of paths $\rho^N(t)$, $t\in[0,T]$ of empirical measures, where $T>0$ is arbitrary, but fixed throughout. As we will explain below, the LDP implies a law of large numbers for $\{\rho^N,\ N\in\mathbb{N}\}$. Both results are shown under mild regularity assumptions on the functions b, σ and on the initial empirical measures $\{\rho^N(0),\ N\in\mathbb{N}\}$. In particular, the stationarity assumption on the process of spacings, which was crucial in the analysis of [37], can be omitted here.

Assumption 1. The function b is Lipschitz on $[0,1]$. Moreover, the deterministic initial empirical measures $\{\rho^N(0),\ N\in\mathbb{N}\}$ converge weakly to ρ_0 as $N\to\infty$ for some probability measure ρ_0 on $\mathbb{R}$, and $\sup_{N\in\mathbb{N}}\int_{\mathbb{R}}|x|^{1+\eta}\,\mathrm{d}\rho^N(0)<\infty$ for some $\eta\in(0,1)$.

Assumption 2. The function $A:=\frac{1}{2}\sigma^2$ is strictly positive, continuously differentiable with a Lipschitz continuous derivative on $[0,1]$. Moreover, the measure ρ_0 has a continuously differentiable density θ with respect to the Lebesgue measure on $\mathbb{R}$ such that $\theta\in L^3(\mathbb{R})$, $\frac{F_{\rho_0}^2}{\theta}\mathbf{1}_{(-\infty,0]}$, $\frac{(1-F_{\rho_0})^2}{\theta}\mathbf{1}_{(0,\infty)}$, $\frac{(\theta')^2}{\theta}\in L^1(\mathbb{R})$ and $\int_{\mathbb{R}}|x|\,\mathrm{d}\rho_0<\infty$, with the convention $\frac{0}{0}=0$ throughout the paper.

We remark that $\frac{F_{\rho_0}^2}{\theta}\mathbf{1}_{(-\infty,0]}$, $\frac{(1-F_{\rho_0})^2}{\theta}\mathbf{1}_{(0,\infty)}\in L^1(\mathbb{R})$ imply, in particular, that the support of ρ_0 is given by a (possibly unbounded) interval, which contains 0. However, one can carry over all our results to measures obtained from ρ_0 by translations, by simply translating the particle systems accordingly.

To state the LDP, we introduce the following notation. We write $M_1(\mathbb{R})$ for the space of probability measures on $\mathbb{R}$, which we view as a metric space with the metric being given by the Lévy distance

(1.5)
$$d_L(\alpha_1,\alpha_2):=\inf\{\epsilon>0|\ \forall\text{ open } O\subset\mathbb{R}:\ \alpha_1(O)\le\alpha_2(O_\epsilon)+\epsilon,\ \alpha_2(O)\le\alpha_1(O_\epsilon)+\epsilon\}.$$

Here, O_ϵ stands for the ϵ-neighborhood of O in $\mathbb{R}$. We let $C([0,T],M_1(\mathbb{R}))$ stand for the space of continuous functions from $[0,T]$ to $M_1(\mathbb{R})$ endowed with the metric

$$d(\gamma_1(\cdot),\gamma_2(\cdot))=\sup_{t\in[0,T]}d_L(\gamma_1(t),\gamma_2(t)),\tag{1.6}$$

and for each $\gamma\in C([0,T],M_1(\mathbb{R}))$ we identify γ with the aggregate cumulative distribution function $R=F_{\gamma(\cdot)}(\cdot)$. We write $\mathbb{R}_T$ for $[0,T]\times\mathbb{R}$, m for the Lebesgue measure on $\mathbb{R}_T$ and $\overline{\mathcal{S}}$ for the space of functions g on $\mathbb{R}_T$, which are infinitely differentiable and such that, for all $t\in[0,T]$, $g(t,\cdot)$ is a Schwartz function on $\mathbb{R}$. Finally, we define $\mathcal{F}$ as the collection of $\gamma\in C([0,T],M_1(\mathbb{R}))$ starting at $\gamma(0)=\rho_0$, such that $t\mapsto\int_{\mathbb{R}}g(t,.)\,\mathrm{d}\gamma(t)$ is absolutely continuous on $[0,T]$ for each $g\in\overline{\mathcal{S}}$, and $R=F_{\gamma(\cdot)}(\cdot)\in C_b(\mathbb{R}_T)$ with

$$R_t,\ R_{xx}\in L^q(\mathbb{R}_T),\ R_x\in L^3(\mathbb{R}_T),\ \int_{\mathbb{R}_T}\frac{R_{xx}^2}{R_x}\,\mathrm{d}m<\infty,\ \int_{\mathbb{R}_T}\frac{R_t^2}{R_x}\,\mathrm{d}m<\infty,\tag{1.7}$$

$$\int_{\mathbb{R}_T}|x|^{1+\eta}R_x\,\mathrm{d}m<\infty,\tag{1.8}$$

for $q=3/2$. Our main result then reads as follows.

Theorem 1.1. *Under Assumptions* 1, 2, *the sequence* $\{\rho^N,\ N\in\mathbb{N}\}$ *satisfies a LDP on* $C([0,T],M_1(\mathbb{R}))$ *with the good rate function*

$$J(\gamma)=\begin{cases}\frac{1}{2}\left\|\frac{R_t-(A(R)R_x)_x+b(R)R_x}{\sigma(R)R_x}\right\|^2_{L^2(\mathbb{R}_T,R_x\,\mathrm{d}m)}, & \textit{if } \gamma\in\mathcal{F}\\ \infty, & \textit{otherwise,}\end{cases}\tag{1.9}$$

and scale N.

In particular, Theorem 1.1 implies the following law of large numbers for the sequence $\{\rho^N,\, N \in \mathbb{N}\}$.

Corollary 1.2. *Under Assumptions 1, 2, the sequence $\{\rho^N,\; N \in \mathbb{N}\}$ converges almost surely to the unique path $\gamma \in \mathcal{F}$, for which the corresponding function R is a generalized solution of the Cauchy problem*

$$R_t = (A(R)R_x)_x - b(R)R_x, \quad R(0,\cdot) = F_{\rho_0}(\cdot)\,. \tag{1.10}$$

Indeed, by [18, Theorem 4] there is a unique nonnegative continuous bounded generalized solution of the problem (1.10) in the sense of [18, Definition 4]. In addition, the explicit formula for the rate function in Theorem 1.1 shows that for every path $\gamma \in \mathcal{F}$, which satisfies $J(\gamma) = 0$, the corresponding function R is a nonnegative continuous bounded generalized solution of the problem (1.10). Therefore, Corollary 1.2 follows from the LDP of Theorem 1.1 and the goodness of the rate function there.

In [8], the authors prove a LDP for systems of diffusions with drift coefficients being continuous functions of the value of the diffusion and the empirical measure of the whole system, and diffusion coefficients being constant and same for all diffusions. There, the LDP is shown by a clever application of the Girsanov Theorem, which allows to move from the system of interacting diffusions to the corresponding system of independent diffusions on the event that the path of empirical measures is near a deterministic path of probability measures and to establish the local LDP. In our case, this approach is not viable due to the interaction through the diffusion coefficients in (1.2). Moreover, the discontinuity of the drift and the diffusion coefficients presents an additional challenge. Even on the level of the law of large numbers as in Corollary 1.2, previous works had to assume that there is no interaction through the diffusions coefficients (see [22], [5], [4] and the references therein), or to work with very special initial conditions (see [37]). We overcome these challenges, but remark that our analysis relies rather heavily on the particular form of the drift and the diffusion coefficients in (1.2).

A crucial part of the proof of Theorem 1.1 is devoted to the study of generalized solutions to porous medium equations with tilt:

$$R_t = (A(R)R_x)_x + h\,A(R)\,R_x. \tag{1.11}$$

The following regularity results, which we need in the proof of Theorem 1.1, are also of independent interest.

Theorem 1.3. *Let $R \in C_b(\mathbb{R}_T)$ be such that, for every $t \in [0,T]$, the function $R(t,\cdot)$ is a cumulative distribution function of a probability measure $\gamma(t)$. Suppose that R is*

a generalized solution to (1.11) *with initial condition* $R(0,\cdot) = F_{\rho_0}(\cdot)$*, where* A *and* ρ_0 *satisfy Assumption* 2 *and* h *is a function on* $\mathbb{R}_T$ *such that*

$$\int_{\mathbb{R}_T} h^2(t,x)\, d\gamma(t)\, \mathrm{d}t < \infty. \tag{1.12}$$

If, in addition, R *satisfies the moment condition*

$$\int_{\mathbb{R}_T} |x|^{1+\iota}\, d\gamma(t)\, \mathrm{d}t < \infty \tag{1.13}$$

for some $\iota \in (0,1)$*, then* (1.7) *holds for all* $\frac{6}{5} \le q \le \frac{3}{2}$*.*

The rest of the paper is devoted to the proofs of Theorems 1.1 and 1.3. In the next section we present the different ingredients of the proofs and explain how they can be combined to obtain the two main theorems.

2. Outline

To be able to describe the outline of the proofs, we need some additional notation. We write (α, f) for $\int_{\mathbb{R}} f\, \mathrm{d}\alpha$ for any f in the space of continuous bounded functions $C_b(\mathbb{R})$ and any $\alpha \in M_1(\mathbb{R})$. Moreover, we let $\tilde{\mathcal{A}}$ be the space of paths $\gamma \in C([0,T], M_1(\mathbb{R}))$ starting at $\gamma(0) = \rho_0$, such that $R = F_{\gamma(t)} \in C_b(\mathbb{R}_T)$ and the moment condition (1.13) holds for $\eta \in (0,1)$ of Assumption 1. Setting $\mathcal{R}_t^\gamma g = g_t + b(R) g_x + A(R) g_{xx}$, the functional

$$\Phi(t,g) = (\gamma(t), g) - (\gamma(0), g) - \int_0^t (\gamma(s), \mathcal{R}_s^\gamma g)\, \mathrm{d}s \tag{2.1}$$

on $\overline{\mathcal{S}}$, with $\Phi(g) := \Phi(T,g)$ and inner product

$$(f,g)_\gamma = \int_{\mathbb{R}_T} f_x\, g_x\, A(R)\, \mathrm{d}\gamma(t)\, \mathrm{d}t \tag{2.2}$$

on $\overline{\mathcal{S}}$, we consider the functional

$$\tilde{I}(\gamma) = \begin{cases} \sup_{g\in\overline{\mathcal{S}}} \big[\Phi(g) - (g,g)_\gamma\big], & \gamma \in \tilde{\mathcal{A}} \\ \infty, & \text{otherwise} \end{cases} \tag{2.3}$$

on $C([0,T], M_1(\mathbb{R}))$. Further, let $\mathcal{A}$ denote the collection of $\gamma \in \tilde{\mathcal{A}}$ such that $t \mapsto (\gamma(t), g(t,\cdot)) : [0,T] \to \mathbb{R}$ is absolutely continuous for each $g \in \overline{\mathcal{S}}$ and consider also the functional

$$I(\gamma) = \begin{cases} \sup_{g\in\overline{\mathcal{S}}} \frac{1}{4} \int_0^T \frac{\left|\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t,g)\right|^2}{(\gamma(t), A(R)(g_x)^2)}\, \mathrm{d}t, & \gamma \in \mathcal{A} \\ \infty, & \text{otherwise} \end{cases} \tag{2.4}$$

on $C([0,T], M_1(\mathbb{R}))$. Note that if $I(\gamma) < \infty$ then for any $g \in \overline{\mathcal{S}}$,

$$\Phi(T,g) = \int_0^T \left[\frac{\mathrm{d}}{\mathrm{d}t}(\gamma(t), g) - (\gamma(t), \mathcal{R}_t^\gamma g)\right] \mathrm{d}t$$

so (taking here $a_2 := (\gamma(t), A(R)(g_x)^2))$, by the elementary inequality $2a_1 - a_2 \le \frac{a_1^2}{a_2}$, $a_1 \in \mathbb{R}$, $a_2 > 0$ we deduce that $\tilde{I}(\gamma) \le I(\gamma) < \infty$.

In addition, we introduce the set $\mathcal{G} \subset \mathcal{F}$ of all paths γ such that $R(t,\cdot) = F_{\gamma(t)}(\cdot)$ is differentiable in t and twice differntiable in x, with R_x strictly positive and (1.11) holding pointwise for some Lipschitz function $h \in C_b(\mathbb{R}_T)$. Finally, for each $\gamma \in C([0,T], M_1(\mathbb{R}))$ and $\delta > 0$, we write $B(\gamma,\delta)$ for the open ball of radius δ around γ in $C([0,T], M_1(\mathbb{R}))$.

As we explain below, Theorems 1.1 and 1.3 are consequences of the following five propositions.

Proposition 2.1. *For all $\gamma \in \mathcal{A}$ with $I(\gamma) < \infty$, one has $J(\gamma) = \tilde{I}(\gamma) \le I(\gamma)$. In particular, the inequality $J(\gamma) \le I(\gamma)$ holds for all $\gamma \in C([0,T], M_1(\mathbb{R}))$.*

Proposition 2.2. *If $\gamma \in \mathcal{A}$ with $\tilde{I}(\gamma) < \infty$ then $R = F_{\gamma(\cdot)}(\cdot)$ satisfies (1.7) for all $\frac{6}{5} \le q \le \frac{3}{2}$. That is:*
(A) $R_x \in L^3(\mathbb{R}_T)$,
(B) R_t, $R_{xx} \in L^q(\mathbb{R}_T)$ for all $\frac{6}{5} \le q \le \frac{3}{2}$,
(C) $\int_{\mathbb{R}_T} \frac{R_{xx}^2}{R_x}\,\mathrm{d}m < \infty$, $\int_{\mathbb{R}_T} \frac{R_t^2}{R_x}\,\mathrm{d}m < \infty$.

Proposition 2.3. *Let $\gamma \in C([0,T], M_1(\mathbb{R}))$ be such that $J(\gamma) < \infty$. Then, there exists a sequence $\{\gamma^M,\ M \in \mathbb{N}\} \subset \mathcal{G}$ such that $\gamma^M \to \gamma$ in $C([0,T], M_1(\mathbb{R}))$ in the limit $M \to \infty$ and*

$$\lim_{M\to\infty} J(\gamma^M) = J(\gamma). \tag{2.5}$$

Proposition 2.4. *For every $\gamma \in C([0,T], M_1(\mathbb{R}))$, the local large deviations upper bound*

$$\lim_{\delta\downarrow 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\rho^N \in B(\gamma,\delta)) \le -I(\gamma) \tag{2.6}$$

holds. Moreover, the sequence $\{\rho^N,\ N \in \mathbb{N}\}$ is exponentially tight in the sense that for every $M > 0$ there exists a compact set $K_M \subset C([0,T], M_1(\mathbb{R}))$, for which

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\rho^N \notin K_M) \le -M. \tag{2.7}$$

Proposition 2.5. *For every $\gamma \in \mathcal{G}$ the local large deviations lower bound*

$$\lim_{\delta\downarrow 0} \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\rho^N \in B(\gamma,\delta)) \ge -J(\gamma) \tag{2.8}$$

holds.

We will show now how Theorems 1.1 and 1.3 can be obtained from the latter five propositions.

Proof of Theorem 1.1. From [10, Theorem 4.1.11] and [10, Lemma 1.2.18] we conclude that our Theorem 1.1 will follow if we can show

$$\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}(\rho^N\in B(\gamma,\delta))\le -J(\gamma)\quad\text{and}\tag{2.9}$$

$$\lim_{\delta\downarrow 0}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}(\rho^N\in B(\gamma,\delta))\ge -J(\gamma)\tag{2.10}$$

for all $\gamma\in C([0,T],M_1(\mathbb{R}))$, and that the sequence $\{\rho^N,\ N\in\mathbb{N}\}$ is exponentially tight in the sense of (2.7). However, (2.9) is a consequence of Propositions 2.4 and 2.1. Moreover, (2.10) holds for all $\gamma\in\mathcal{G}$ by Proposition 2.5. Now, for every fixed $\gamma\in\mathcal{F}$ and $\delta>0$, we can pick a sequence $\{\gamma^M\}\subset\mathcal{G}$ as in Proposition 2.3 such that $B(\gamma^M,\delta'/2)\subset B(\gamma,\delta)$ for all $\delta'<\delta$ and $M\ge M_0(\delta)$. Taking the limit $\delta'\downarrow 0$ followed by $M\uparrow\infty$, we conclude

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}(\rho^N\in B(\gamma,\delta))\ge -\liminf_{M\to\infty}J(\gamma^M).\tag{2.11}$$

Passing to the limit $\delta\downarrow 0$ and using Proposition 2.3, one ends up with (2.10) for all $\gamma\in\mathcal{F}$. Moreover, for all $\gamma\notin\mathcal{F}$, the bound (2.10) is trivially true. Lastly, exponential tightness of the sequence $\{\rho^N,\ N\in\mathbb{N}\}$ is part of Proposition 2.4. $\square$

Proof of Theorem 1.3. We first note after integration by parts in space, that having $R=F_{\gamma(\cdot)}(\cdot)$ as generalized solution of (1.11), is equivalent to

$$\Phi(t,g)=-\int_0^t(\gamma(s),b(R)g_x+hA(R)g_x)\mathrm{d}s\,,\tag{2.12}$$

for any $g\in\mathcal{S}$ and $t\in[0,T]$. Thus, taking here $b\equiv 0$ without loss of generality, and denoting by $\overline{\mathcal{S}}_x$ the space of spatial derivatives $f=g_x$ of $g\in\overline{\mathcal{S}}$, we see that for γ as in Theorem 1.3,

$$\begin{aligned}\tilde{I}(\gamma)&=\sup_{f\in\overline{\mathcal{S}}_x}\Big[\int_{\mathbb{R}_T}\big(-h\,A(R)f-A(R)f^2\big)\,\mathrm{d}\gamma(t)\,\mathrm{d}t\Big]\\&\le\sup_{f\in L^2(\mathbb{R}_T,\,\mathrm{d}\gamma(t)\mathrm{d}t)}\Big[\int_{\mathbb{R}_T}\big(-h\,A(R)f-A(R)f^2\big)\,\mathrm{d}\gamma(t)\,\mathrm{d}t\Big].\end{aligned}$$

The latter supremum is attained for $f=-\frac{1}{2}h$ and its value is finite due to our assumption (1.12). Consequently, in this case $\tilde{I}(\gamma)<\infty$. Our assumptions in Theorem 1.3 further guarantee that $\gamma\in\mathcal{A}$ (with $\eta=\iota$), hence in view of Proposition 2.2, such $R=F_{\gamma(t)}$ satisfies the regularity properties (1.7), as claimed. $\square$

In the following six sections we give the proofs of the propositions above in the order in which they are stated. The proofs of the Propositions 2.1, 2.2 and 2.3 are analytic for the most part and rely on results from [25, 26, 28] concerning linear and non-linear parabolic equations with non-smooth coefficients. The proofs of Propositions 2.4 and 2.5 are probabilistic for the most part and involve tools from the theory of large deviations and stochastic analysis.

3. Proof of Proposition 2.1

We fix $\gamma \in \mathcal{A}$ with $I(\gamma) < \infty$, let $R = F_{\gamma(.)}(.)$ and consider the Hilbert space $\mathbb{H}$ given by the closure of $\overline{\mathcal{S}}$ under the norm corresponding to the inner product $(\cdot,\cdot)_\gamma$ of (2.2). As remarked in the outline, $\tilde{I}(\gamma) \le I(\gamma) < \infty$, so that

$$\sup_{g\in\overline{\mathcal{S}}} \big[\Phi(g) - (g,g)_\gamma\big] = \tilde{I}(\gamma) < \infty.$$

Thus, Φ is a bounded linear functional on $\overline{\mathcal{S}}$ with respect to the norm induced by the scalar product $(\cdot,\cdot)_\gamma$. Consequently, there exists a unique bounded linear functional $\overline{\Phi}$ on $\mathbb{H}$, whose restriction to $\overline{\mathcal{S}}$ coincides with Φ. Now, by the Riesz representation theorem, there is a unique element $\tilde{h} \in \mathbb{H}$, which satisfies $\overline{\Phi}(g) = (\tilde{h}, g)_\gamma$ for all $g \in \mathbb{H}$. Combining this with the fact that $\overline{\mathcal{S}}$ is by definition dense in $\mathbb{H}$, we obtain

$$\tilde{I}(\gamma) = \sup_{g\in\mathbb{H}} \big[\overline{\Phi}(g) - (g,g)_\gamma\big] = \sup_{g\in\mathbb{H}} \big[(\tilde{h},g)_\gamma - (g,g)_\gamma\big] = \frac{1}{4}(\tilde{h},\tilde{h})_\gamma. \tag{3.1}$$

Furthermore, by the definition of $\tilde{h}$ and Φ, $\tilde{h}_x \in L^2(\mathbb{R}_T, \mathrm{d}\gamma(t)\,\mathrm{d}t)$ satisfies

$$\Phi(t,g) = \int_0^t (\gamma(s), \tilde{h}_x\, g_x\, A(R))\,\mathrm{d}s \tag{3.2}$$

for $t = T$ and any $g \in \overline{\mathcal{S}}$. In particular, considering Schwartz functions g supported on $\mathbb{R}_t$ we have that (3.2) applies also for any $t \in [0,T]$. Comparing this with (2.12) we deduce that R is a generalized solution of the partial differential equation (1.11) for

$$h = -\tilde{h}_x - \frac{b(R)}{A(R)}\,. \tag{3.3}$$

In view of the boundedness of $\frac{b}{A}$ (due to Assumptions 1 and 2), we have that $h \in L^2(\mathbb{R}_T, \mathrm{d}\gamma(t)\,\mathrm{d}t)$. By Theorem 1.3 this implies that R_t, R_{xx} and the $L^1(\mathbb{R}_T)$ density R_x are elements of $L^{3/2}(\mathbb{R}_T)$ and moreover, the functions $\frac{R_t}{R_x}$, R_x and $\frac{R_{xx}}{R_x}$ are elements of $L^2(\mathbb{R}_T, R_x\,\mathrm{d}m)$. Thus, the identity

$$h = \frac{R_t - (A(R)R_x)_x}{A(R)R_x} \tag{3.4}$$

holds in $L^2(\mathbb{R}_T, R_x\,\mathrm{d}m)$. Finally, putting (3.4), (3.3) and (3.1) together, we end up with

$$I(\gamma) \ge \tilde{I}(\gamma) = \frac{1}{4}\left\|\frac{R_t - (A(R)R_x)_x + b(R)R_x}{A(R)^{1/2}R_x}\right\|^2_{L^2(\mathbb{R}_T, R_x\,\mathrm{d}m)} = J(\gamma), \tag{3.5}$$

as stated. $\square$

4. Proof of Proposition 2.2 (a) and (b)

Fixing $R = F_\gamma$ for $\gamma \in \mathcal{A}$ with $\tilde{I}(\gamma) < \infty$, we prove Proposition 2.2 (A) in a series of four lemmas, starting with a-priori local L^2 estimate on R_x. Combining it in Lemma 4.2 with the moment condition (1.13) we deduce a global $L^{3/2}$ estimate on R_x, which we improve in 4.3 to L^p estimates on R_x for all $\frac{3}{2} \le p < 3$. Finally,

we prove in Lemma 4.4 uniform boundedness of the corresponding norms, resulting with $R_x \in L^3(\mathbb{R}_T)$.

Lemma 4.1. *If $\gamma \in \mathcal{A}$ is such that $I(\gamma) < \infty$, then:*

(a) The measure $\mathrm{d}\gamma(t)\,\mathrm{d}t$ on $\mathbb{R}_T$ has a density with respect to the Lebesgue measure on $\mathbb{R}_T$, whose L^2 norm restricted to any strip $S_n = [0,T] \times [n, n+1]$ is bounded by a constant $C(T, I(\gamma)) < \infty$, which does not depend on the value of $n \in \mathbb{Z}$. In particular, this density is locally square integrable, so that the weak derivative R_x exists as a function in $L^2_{loc}(\mathbb{R}_T)$.

(b) The function $a := A(R)$ is uniformly continuous on $\mathbb{R}_T$.

Proof. To prove part (a) of the lemma, it suffices to show that there is a constant $C(T, I(\gamma)) < \infty$ such that the inequality

$$\int_{S_n} \psi(t,x)\,\mathrm{d}\gamma(t)\,\mathrm{d}t \leq C(T, I(\gamma))\|\psi\|_{L^2(S_n)}, \tag{4.1}$$

holds for all $n \in \mathbb{Z}$ and all non-negative continuos functions $\psi : \mathbb{R}_T \to \mathbb{R}$ supported on the strip S_n. To this end, we use the Cauchy-Schwarz inequality and the definition of I to conclude

$$\begin{aligned} &\int_{\mathbb{R}} g(T,\cdot)\,\mathrm{d}\gamma(T) - \int_{\mathbb{R}} g(0,\cdot)\,\mathrm{d}\gamma(0) + \int_{\mathbb{R}_T} (g_t + b(R)g_x + a g_{xx})\,\mathrm{d}\gamma(t)\,\mathrm{d}t \\ &\qquad \leq 2\,I(\gamma)^{1/2}\left(\int_{\mathbb{R}_T} (g_x)^2 a(t,x)\,\mathrm{d}\gamma(t)\,\mathrm{d}t\right)^{1/2} \end{aligned} \tag{4.2}$$

for all $g \in \overline{\mathcal{S}}$.

Fixing first a strip $S = [0,T] \times [-1/2, 1/2]$, we shall use [25, Theorem 2], which in case $d = 1$ provides space-time smoothing kernels $k^\epsilon(t,x) = \epsilon^{-2}\xi(t/\epsilon)\xi(x/\epsilon)$, $\epsilon > 0$, for some infinitely differentiable, probability density function $\xi(\cdot)$ of compact support, such that for ψ as above, there exists a bounded measurable function $z : \mathbb{R}_T \to \mathbb{R}$, which is non-positive almost everywhere, convex in x on $[0,T] \times [-1,1]$ and zero outside of $[0,T] \times [-1,1]$, increasing in t and for any $\epsilon > 0$ small enough, the functions $\psi^\epsilon = \psi * k^\epsilon$ and $z^\epsilon = z * k^\epsilon$ satisfy on S the following inequalities:

$$\forall c \geq 0: \quad c^{1/2}\psi^\epsilon \leq C\left(z^\epsilon_t + c\, z^\epsilon_{xx}\right), \tag{4.3}$$

$$\frac{1}{2}|z^\epsilon_x| \leq -z^\epsilon \leq C\,\|\psi\|_2\,, \tag{4.4}$$

with $C > 0$ a universal constant independent of ψ and ϵ. In the preceeding the compact support of z is specified in the proof of [25, Theorem 2] after display (29).

Next, we use first (4.3) with $c = \inf_{\mathbb{R}_T} a$, then the non-negativity of $z^\epsilon_t = z_t * k^\epsilon$ and $z^\epsilon_{xx} = z_{xx} * k^\epsilon$, followed by the bound (4.2) for z^ϵ and Jensen's inequality, and

finally applying (4.4), we arrive at the bound:

$$\begin{aligned}
&c^{1/2}\int_S \psi^\epsilon \,\mathrm{d}\gamma(t)\,\mathrm{d}t \le C\int_S (z^\epsilon_t + a\, z^\epsilon_{xx})\,\mathrm{d}\gamma(t)\,\mathrm{d}t\\
\le\;& C\int_{\mathbb{R}_T} (z^\epsilon_t + a\, z^\epsilon_{xx})\,\mathrm{d}\gamma(t)\,\mathrm{d}t\\
\le\;& \big(2\,I(\gamma)^{1/2} + \|b\|_\infty T^{1/2}c^{-1/2}\big)\Big(\int_{\mathbb{R}_T}(z^\epsilon_x)^2 a\,\mathrm{d}\gamma(t)\,\mathrm{d}t\Big)^{1/2} - \int_{\mathbb{R}} z^\epsilon(T,\cdot)\,\mathrm{d}\gamma(T)\\
\le\;& C\,\|\psi\|_2\big[2\big(2\,I(\gamma)^{1/2} + \|b\|_\infty T^{1/2}c^{-1/2}\big)\|a\|_\infty^{1/2}\,T^{1/2}+1\big].
\end{aligned}$$

Taking the limit $\epsilon \downarrow 0$, we end up with (4.1) with S_n replaced by S and some constant $C(T, I(\gamma))$, which depends only on T, $I(\gamma)$, the functions a, b and the dimension of the problem.

To complete the proof, we can consider the same problem with ρ_0 replaced by $\rho_0(\cdot + n + 1/2)$ and I replaced by the corresponding rate function I_n. Then, for the path of measures $\gamma_n(t)(\cdot) = \gamma(t)(\cdot + n + 1/2)$, $t \in [0,T]$ it holds $I_n(\gamma_n) = I(\gamma)$. Further, to any continuous non-negative $\psi : \mathbb{R}_T \to \mathbb{R}$ supported on S_n corresponds $\psi_n(\cdot,\cdot) = \psi(\cdot, \cdot + n + 1/2)$ supported on S, such that by the preceeding proof:

$$\int_{S_n} \psi(t,x)\,\mathrm{d}\gamma(t)\,\mathrm{d}t = \int_S \psi_n(t,x)\;\mathrm{d}\gamma_n(t)\,\mathrm{d}t \le C(T, I(\gamma))\|\psi_n\|_2 = C(T, I(\gamma))\|\psi\|_2\,.$$

This finishes the proof of part (a) of the lemma.

Since the function A is assumed to be Lipschitz, it suffices to show that R is uniformly continuous on $\mathbb{R}_T$. However, we know that the assumption of finite rate of γ implies that the function R is continuous, thus, uniformly continuous on compact sets. In addition, the continuity of $t \mapsto \gamma(t)$ with respect to the weak topology shows that the set $\{\gamma(t)\}_{t\in[0,T]}$ is compact and by Prokhorov's Theorem uniformly tight. Hence, for every $\epsilon > 0$, there is an $M > 0$ such that

$$\sup_{t\in[0,T],x\ge M} \max(R(t,-x), 1 - R(t,x)) < \epsilon. \tag{4.5}$$

The latter two observations show that R is uniformly continuous on $\mathbb{R}_T$. □

We will now establish Proposition 2.2 (A) in a series of three lemmas.

Lemma 4.2. *Let $\gamma \in \mathcal{A}$ be such that $I(\gamma) < \infty$, and set $R = F_{\gamma(\cdot)}(\cdot)$ as before. Then, R_x exists as a function in $L^{3/2}(\mathbb{R}_T)$.*

Proof. We start by fixing an $n \in \mathbb{Z}$ and applying the Cauchy-Schwarz inequality with respect to the Lebesgue measure on S_n to obtain

$$\int_{S_n} R_x^{1/2}\,R_x\,\mathrm{d}m \le \Big(\int_{S_n} R_x\,\mathrm{d}m\Big)^{1/2}\Big(\int_{S_n} R_x^2\,\mathrm{d}m\Big)^{1/2}. \tag{4.6}$$

Since our goal is to prove that the left side is summable over all $n \in \mathbb{Z}$ and we know from the Lemma 4.1 that the second factor on the right side is bounded above by a uniform constant, it remains to show that

$$\sum_{n\in\mathbb{Z}} \Big(\int_{S_n} R_x \,\mathrm{d}m \Big)^{1/2} < \infty. \tag{4.7}$$

We will only prove that $\sum_{n\in\mathbb{Z}_-} \big(\int_{S_n} R_x \,\mathrm{d}m \big)^{1/2} < \infty$, since the sum over all $n \in \mathbb{Z}_+$ can be dealt with in the same way. To this end, we apply the Cauchy-Schwarz inequality to conclude

$$\begin{aligned} \sum_{n\in\mathbb{Z}_-} \Big(\int_{S_n} R_x \,\mathrm{d}m \Big)^{1/2} |n|^{1+\eta} |n|^{-1-\eta} &\le C_1 \sum_{n\in\mathbb{Z}_-} \Big(\int_{S_n} R_x \,\mathrm{d}m \Big) |n|^{1+\eta} \\ \le\; C_1 \int_{S_{-1}} R_x \,\mathrm{d}m + C_2 \Big(\sum_{n\le -2} \int_{S_n} R_x(t,y) |y|^{1+\eta} \,\mathrm{d}m(t,y) \Big) & \\ =\; C_1 \int_{S_{-1}} R_x \,\mathrm{d}m + C_2 \int_{[0,T]\times(-\infty,-1]} R_x(t,y) |y|^{1+\eta} \,\mathrm{d}m(t,y) < \infty & \end{aligned}$$

for some uniform constants $C_1, C_2 > 0$ and with η of Assumption 1. The lemma now readily follows. $\square$

Lemma 4.3. *Let $\gamma \in \mathcal{A}$ be such that $I(\gamma) < \infty$ and write $R = F_{\gamma(\cdot)}(\cdot)$. Then, R_x exists as a function in $L^p(\mathbb{R}_T)$ for all $\frac{3}{2} \le p < 3$.*

Proof. We fix a $\frac{3}{2} \le p < 3$ and define $\frac{3}{2} < q \le 3$ by $\frac{1}{p} + \frac{1}{q} = 1$. Now, we pick an $f \in \overline{\mathcal{S}}$ taking only nonnegative values and consider the backward Cauchy problem

$$u_t + a(t,x)\, u_{xx} + f(t,x)\, u = 0, \quad u(T,\cdot) = 1 \tag{4.8}$$

with $a = A(R)$ as before. Noting that $f \in L^q(\mathbb{R}_T) \cap L^2(\mathbb{R}_T)$ and that the norm bounds of [26, Theorem 4.1] can be refined to an estimate on the norms in $L^q(\mathbb{R}_T) \cap L^2(\mathbb{R}_T) \cap L^3(\mathbb{R}_T)$, we apply the method of continuity (see e.g. [29, section III.2]) to conclude that this problem has a weak solution, for which $v := u - 1$ is an element of $W^{1,2}_q(\mathbb{R}_T) \cap W^{1,2}_2(\mathbb{R}_T) \cap W^{1,2}_3(\mathbb{R}_T)$. Next, we let P be the law of a diffusion with generator $a(t,x)\, \frac{\mathrm{d}^2}{\mathrm{d}x^2}$ and apply Itô's formula in the form of the last identity in [24, chapter 10, Theorem 1] to obtain the stochastic representation

$$u(t,x) = \mathbb{E}^P \Big[\exp\Big(\int_t^T f(s, x(s))\,\mathrm{d}s \Big) \Big| x(t) = x \Big], \tag{4.9}$$

where $x(\cdot)$ is the canonical process. It follows immediately that $u \ge 1$, and by Portenko's Lemma (see [34, inequality (6)] and note that it only relies on the standard heat kernel estimate for the diffusion with law P), there is a non-decreasing function $G_q :\, (0,\infty) \to (0,\infty)$ depending only on $\frac{3}{2} < q \le 3$ (but not on f) such that

$$|u(t,x)| \le G_q\big(\|f\|_{L^q(\mathbb{R}_T)} \big), \quad (t,x) \in \mathbb{R}_T. \tag{4.10}$$

Next, we observe that the function $v := \log u$ is a generalized solution of the problem

$$v_t + a(t,x)\,v_{xx} + a(t,x)\,v_x^2 + f(t,x) = 0, \quad v(T,\cdot) = 0 \tag{4.11}$$

and inherits the bound

$$|v(t,x)| \le G_q\big(\|f\|_{L^q(\mathbb{R}_T)}\big), \quad (t,x) \in \mathbb{R}_T, \tag{4.12}$$

where we have increased G_q if necessary. We claim the following chain of estimates:

$$\begin{aligned}(\gamma(T), v(T,\cdot)) - (\gamma(0), v(0,\cdot)) &= \int_{\mathbb{R}_T} (v_t + a\,v_{xx} + v_x\,h\,a)\,R_x\,\mathrm{d}m \\ &= \int_{\mathbb{R}_T} (-a\,v_x^2 + v_x\,h\,a - f)\,R_x\,\mathrm{d}m \\ &\le \int_0^T \int_{\mathbb{R}} \Big(\frac{1}{4}h^2\,a - f\Big)\,R_x\,\mathrm{d}m.\end{aligned}$$

Here, the first identity is a consequence of the definition of h (see (3.2) and (3.3)) and the fact that v is bounded, $v_t,\ v_{xx} \in L^3(\mathbb{R}_T)$, $v_x \in L^6(\mathbb{R}_T)$ (by the parabolic Sobolev inequality in the form of [28, chapter II, Lemma 3.3]), which allows to approximate v in the corresponding norm by functions in $\overline{\mathcal{S}}$; the second identity follows from (4.11); and the inequality can be obtained by optimizing the expression $(-a\,v_x^2 + v_x\,h\,a)$ pointwise. Rearranging the terms and using (3.1), we have

$$\int_{\mathbb{R}_T} f\,R_x\,\mathrm{d}m \le \sup_{(t,x)\in\mathbb{R}_T} |v(t,x)| + (C_1 I(\gamma) + C_2), \tag{4.13}$$

where the constants $C_1,\ C_2 < \infty$ depend only on the supremum and the infimum of A, the supremum of $|b|$ and on T. Putting this together with (4.12), we end up with $R_x \in L^p(\mathbb{R}_T)$ as desired. □

Lemma 4.4. *Let $\gamma \in \mathcal{A}$ be such that $I(\gamma) < \infty$ and write $R = F_{\gamma(\cdot)}(\cdot)$. Then, R_x exists as a function in $L^3(\mathbb{R}_T)$.*

Proof. Step 1. In addition to the Cauchy problem

$$R_t - (A(R)R_x)_x = h\,A(R)R_x\,, \quad R(0,\cdot) = F_{\rho_0} \tag{4.14}$$

for R, we consider the solutions to the following two auxilliary Cauchy problems:

$$V_t - V_{xx} = h\,A(R)R_x\,, \quad V(0,\cdot) = 0, \tag{4.15}$$

$$Z_t - Z_{xx} = 0\,, \quad Z(0,\cdot) = F_{\rho_0}, \tag{4.16}$$

obtained by convolving in space with the heat kernel $p(t,x) = (4\pi t)^{-1/2}\exp(-x^2/4t)$. Namely,

$$\begin{aligned}V(t,x) &= \int_0^t \int_{\mathbb{R}} (h\,A(R)R_x)(s,y)p(t-s,x-y)\,\mathrm{d}y\,\mathrm{d}s, \\ Z(t,x) &= \int_{\mathbb{R}} F_{\rho_0}(y)p(t,x-y)\,\mathrm{d}y.\end{aligned}$$

A direct computation shows that $U = R - V - Z$ then solves the Cauchy problem

$$U_t - (A(R)U_x)_x = \big[(A(R)-1)V_x + (A(R)-1)Z_x\big]_x, \quad U(0,\cdot) = 0. \tag{4.17}$$

Our general strategy is to obtain integrability estimates on V_x, Z_x and U_x, and then to deduce the integrability estimate on R_x from those.

To start with, we apply Hölder's inequality in the form of

$$\int_{\mathbb{R}_T} |h|^q\, f^q\, \mathrm{d}m \le \Big(\int_{\mathbb{R}_T} |h|^2\, f\, \mathrm{d}m\Big)^{q/2} \Big(\int_{\mathbb{R}_T} f^p\, \mathrm{d}m\Big)^{(q-1)/(p-1)} \tag{4.18}$$

with a $\frac{3}{2} \le p < 3$ and $q := \frac{2p}{p+1} \in [6/5, 3/2)$ to deduce $h\, A(R) R_x \in L^q(\mathbb{R}_T)$. Thus, by the regularity theory for the heat equation (see inequalities (3.1) and (3.2) in [28, chapter IV], or [26, Theorem 2.1]), we have

$$\|V\|_{W^{1,2}_q(\mathbb{R}_T)} \le C_1 \|h\, A(R) R_x\|_{L^q(\mathbb{R}_T)} \tag{4.19}$$

where $C_1 < \infty$ is a uniform constant (which in particular does not depend on q as long as q belongs to a compact interval). We conclude $V \in W^{1,2}_q(\mathbb{R}_T)$ and, thus, V, $V_x \in L^{p'}(\mathbb{R}_T)$ with $p' = \big(\frac{1}{q} - \frac{1}{3}\big)^{-1} = \frac{6p}{3+p}$ due to the parabolic Sobolev inequality in the form of [28, chapter II, Lemma 3.3], in which the constants can be chosen uniformly for all p in any given compact set and the corresponding $q = \frac{2p}{p+1}$. Moreover, since $\theta \in L^1(\mathbb{R}) \cap L^3(\mathbb{R})$ by Assumption 2, the norms $\|\theta\|_{L^p(\mathbb{R})}$, $1 \le p \le 3$ can be bounded above by a uniform constant. In addition, we note that Z_x is given by the convolution of θ with the standard heat kernel. Putting these two observations together with Young's inequality and Fubini's Theorem, we conclude that $Z_x \in L^p(\mathbb{R}_T)$ for all $1 \le p \le 3$ and that the norms $\|Z_x\|_{L^p(\mathbb{R}_T)}$, $1 \le p \le 3$ can be bounded above by a uniform constant.

Step 2. We observe next that, refining the norm estimates in [26, Theorem 6.2] to estimates in $L^p(\mathbb{R}_T) \cap L^{p'}(\mathbb{R}_T)$ with $p' = \frac{6p}{3+p}$ and applying the method of continuity (see e.g. [29, section III.2]), we may deduce the existence of a solution $\hat{U}$ of the problem (4.17) in the space $\mathcal{H}_p(T) \cap \mathcal{H}_{p'}(T)$ defined in [26]. In particular, $\hat{U} \in L^{p'}(\mathbb{R}_T)$, $\hat{U}_x \in L^{p'}(\mathbb{R}_T)$ and

$$\|\hat{U}\|_{L^{p'}(\mathbb{R}_T)} + \|\hat{U}_x\|_{L^{p'}(\mathbb{R}_T)} \le C_2\big(\|V_x\|_{L^{p'}(\mathbb{R}_T)} + \|Z_x\|_{L^{p'}(\mathbb{R}_T)}\big) \tag{4.20}$$

with $p' = \frac{6p}{3+p}$ and where the constant $C_2 < \infty$ can be chosen uniformly for all $2 \le p' < 3$. We claim now that $U = \hat{U}$ Lebesgue almost everywhere.

Assuming this claim, we conclude that $U_x \in L^{p'}(\mathbb{R}_T)$. Consequently, we also have $R_x = U_x + V_x + Z_x \in L^{p'}(\mathbb{R}_T)$. Moreover, combining the norm bounds with Hölder's inequality (4.18), we get the estimate

$$\|R_x\|_{L^{p'}(\mathbb{R}_T)} \le C_3\, \|R_x\|_{L^p(\mathbb{R}_T)}^{\frac{p(q-1)}{q(p-1)}} + C_4 \tag{4.21}$$

with some constants C_3, $C_4 < \infty$ independent of $\frac{3}{2} \le p < 3$. Therefore, by Jensen's inequality with respect to the measure $R_x \,\mathrm{d}m$ on $\mathbb{R}_T$, it holds

$$\|R_x\|_{L^p(\mathbb{R}_T)}^{\frac{p'(p-1)}{(p'-1)p}} \le C_5 \|R_x\|_{L^p(\mathbb{R}_T)}^{\frac{p(q-1)}{q(p-1)}} + C_6 \tag{4.22}$$

with some constants C_5, $C_6 < \infty$ independent of $\frac{3}{2} \le p < 3$. Finally, we note that in the limit $p \uparrow 3$ the exponent on the left side of (4.22) tends to 1, whereas the exponent on the right side of (4.22) tends to $\frac{1}{2}$. This shows that the norms $\|R_x\|_{L^p(\mathbb{R}_T)}$, $\frac{3}{2} \le p < 3$ can be bounded by a constant independent of p, and we end up with $R_x \in L^3(\mathbb{R}_T)$ as desired.

Step 3. It remains to show that $U = \hat{U}$ Lebesgue almost everywhere. To this end, we prove first that $U \in L^p(\mathbb{R}_T)$ and $U_x \in L^p(\mathbb{R}_T)$. Since we already know that $V \in L^p(\mathbb{R}_T)$, for the first inclusion it suffices to show that $R - Z \in L^p(\mathbb{R}_T)$. To this end, we set $\mathbb{R}_T^+ = [0,T] \times \mathbb{R}_+$, $\mathbb{R}_T^- = [0,T] \times \mathbb{R}_-$ and make the decomposition

$$\int_{\mathbb{R}_T} |R - Z|^p \,\mathrm{d}m = \int_{\mathbb{R}_T^+} |(1-R) - (1-Z)|^p \,\mathrm{d}m + \int_{\mathbb{R}_T^-} |R - Z|^p \,\mathrm{d}m. \tag{4.23}$$

Using the elementary inequality

$$|x_1 - x_2|^p \le 2^{p-1}(|x_1|^p + |x_2|^p), \quad x_1, x_2 \in \mathbb{R},\ p > 1 \tag{4.24}$$

with our p, we can bound the latter expression by a constant multiple of

$$\int_{\mathbb{R}_T^+} (1-R)^p \,\mathrm{d}m + \int_{\mathbb{R}_T^+} (1-Z)^p \,\mathrm{d}m + \int_{\mathbb{R}_T^-} R^p \,\mathrm{d}m + \int_{\mathbb{R}_T^-} Z^p \,\mathrm{d}m.$$

Since $0 \le R \le 1$ and $0 \le Z \le 1$, we may remove the power p to obtain an upper bound on the latter expression. The resulting quantity is then finite if and only if the quantity

$$\int_{\mathbb{R}_T} R_x |x| \,\mathrm{d}m + \int_{\mathbb{R}_T} Z_x |x| \,\mathrm{d}m$$

is finite, and in this case the two quantities are equal. However, the first summand in the latter expression is finite due to $\gamma \in \mathcal{A}$, and the second summand is at most $\frac{2\sqrt{2}}{3} T^{3/2} + T \int_{\mathbb{R}} |x| \,\mathrm{d}\rho_0$ due to the representation of Z_x as the collection of densities of a suitable Brownian motion. All in all, we end up with the inclusion $U \in L^p(\mathbb{R}_T)$. Next, we note that $U_x \in L^p(\mathbb{R}_T)$ follows from $U_x = R_x - V_x - Z_x$, $R_x \in L^p(\mathbb{R}_T)$, $V_x \in L^p(\mathbb{R}_T)$ (due to $V_x \in L^{p'}(\mathbb{R}_T) \cap L^q(\mathbb{R}_T)$) and $Z_x \in L^p(\mathbb{R}_T)$ (see step 2). Here, we have set $p' = \frac{6p}{3+p}$ and $q = \frac{2p}{p+1}$ as before.

Next, we pick a test function $g \in \overline{\mathcal{S}}$ and define the function $f : \mathbb{R}_T \to \mathbb{R}$ by

$$f(t,x) = -\int_x^\infty g(t,y) \,\mathrm{d}y, \quad (t,x) \in \mathbb{R}_T. \tag{4.25}$$

We claim that, for Lebesgue almost every $0 \le t_1 < t_2 \le T$,

$$\begin{aligned}(4.26)\qquad &\int_{\mathbb{R}} f(t_2,x)\,U(t_2,x)\,\mathrm{d}x - \int_{\mathbb{R}} f(t_1,x)\,U(t_1,x)\,\mathrm{d}x - \int_{t_1}^{t_2}\int_{\mathbb{R}} U\,f_t\,\mathrm{d}m \\ &\qquad = -\int_{t_1}^{t_2}\int_{\mathbb{R}} A(R)\,f_x\,U_x - \big[(A(R)-1)(V_x+Z_x)\big]_x\,f\,\mathrm{d}m.\end{aligned}$$

Indeed, one may deduce this from the weak formulation of the PDE in (4.17) via an approximation of f by functions in $\overline{\mathcal{S}}$ using the boundedness of f and $f_t,\ f_x \in \overline{\mathcal{S}}$ and $U \in L^1(\mathbb{R}_T)$ (the latter can be shown in the same way as $U \in L^p(\mathbb{R}_T)$). Now, integrating by parts in space on the left side of (4.26), we conclude that the function $W(t,x) = \int_{-\infty}^x U(t,y)\,\mathrm{d}y$ is a generalized solution of the Cauchy problem

$$(4.27)\qquad W_t = A(R)W_{xx} + (A(R)-1)(V_x+Z_x), \quad W(0,\cdot) = 0$$

on $\mathbb{R}_T$. This, $W_{xx} = U_x \in L^p(\mathbb{R}_T)$, $V_x \in L^p(\mathbb{R}_T)$ and $Z_x \in L^p(\mathbb{R}_T)$ imply $W_t \in L^p(\mathbb{R}_T)$. Thus, in view of [26, norm estimate (6.1)], we have $U \in \mathcal{H}_p(\mathbb{R}_T)$, so that $U = \hat{U}$ Lebesgue almost everywhere by [26, Theorem 2.4]. $\square$

We now turn to the proof of Proposition 2.2 (B).

Proof of Proposition 2.2 (B). We will only show that $R_t,\ R_{xx} \in L^{3/2}(\mathbb{R}_T)$, since the other integrability estimates can be obtained in an analogous manner. We claim first that $U_t,\ U_{xx} \in L^{3/2}(\mathbb{R}_T)$. Indeed, we can rewrite the PDE for U as

$$(4.28)\qquad U_t - A(R)U_{xx} = A'(R)R_x\,U_x + \big[(A(R)-1)V_x + (A(R)-1)Z_x\big]_x.$$

Moreover, since $R_x \in L^3(\mathbb{R}_T)$, $U_x \in L^3(\mathbb{R}_T)$, $V_x \in L^3(\mathbb{R}_T)$, $Z_x \in L^3(\mathbb{R}_T)$, $V_{xx} \in L^{3/2}(\mathbb{R}_T)$, $Z_{xx} \in L^{3/2}(\mathbb{R}_T)$ and the functions A and A' are bounded, the right side in (4.28) belongs to $L^{3/2}(\mathbb{R}_T)$. Thus, we may apply [26, Theorem 2.1] to deduce that there is a function $\tilde{U} \in W^{1,2}_{3/2}(\mathbb{R}_T)$, which satisfies

$$(4.29)\qquad \tilde{U}_t - A(R)\tilde{U}_{xx} = A'(R)R_x\,U_x + \big[(A(R)-1)V_x + (A(R)-1)Z_x\big]_x$$

and $\tilde{U}(0,.) = 0$. Now, we let ϕ_k, $k \in \mathbb{N}$ be a truncation sequence such that $\phi_k \in C^\infty(\mathbb{R})$, $0 \le \phi_k \le 1$, $\phi_k \equiv 1$ on $[-k,k]$, $\phi_k \equiv 0$ on $(-\infty,-k-1] \cup [k+1,\infty)$ and $\max(|\phi_k'|,|\phi_k''|) \le 2$. Next, we fix a $k \in \mathbb{N}$ and set $\tilde{\Delta} = \phi_k(U-\tilde{U})$. Then, $\tilde{\Delta}$ is a generalized solution of the problem

$$(4.30)\qquad \tilde{\Delta}_t - (A(R)\tilde{\Delta}_x)_x + A'(R)R_x\tilde{\Delta}_x = \tilde{\psi}_k, \quad \Delta(0,.) = 0,$$

where

$$(4.31)\qquad \tilde{\psi}_k = -A(R)\phi_k''(U-\tilde{U}) - 2A(R)\phi_k'(U-\tilde{U})_x.$$

Now, a careful reading of the proof of [28, chapter III, Theorem 3.3] shows that the solution of the problem (4.30) in the space $W_2^{0,1}(\mathbb{R}_T)$ is unique and satisfies

$$(4.32)\qquad \|\tilde{\Delta}\|_{W_2^{0,1}(\mathbb{R}_T)} \le C_9\Big(\int_0^T \Big(\int_{\mathbb{R}} \tilde{\psi}_k^{\,2r_1}\,\mathrm{d}x\Big)^{r_2/r_1}\mathrm{d}t\Big)^{1/(2r_2)}$$

for all $r_1 \in [1,\infty]$, $r_2 \in [1,2]$ with $\frac{1}{2r_1} + \frac{1}{r_2} = 1$, provided that

$$\int_0^T \Big(\int_{\mathbb{R}} (A'(R)R_x)^{2r_1}\,\mathrm{d}x \Big)^{r_2/r_1}\,\mathrm{d}t < \infty. \tag{4.33}$$

We choose $r_1 = r_2 = \frac{3}{2}$, so that the latter condition is satisfied. In addition, with this choice, we have

$$\int_0^T \Big(\int_{\mathbb{R}} \tilde{\psi}_k^{\,2r_1}\,\mathrm{d}x \Big)^{r_2/r_1}\,\mathrm{d}t \to 0,\ k \to \infty \tag{4.34}$$

due to $U,\, U_x \in L^3(\mathbb{R}_T)$ (see steps 2 and 3), $\tilde{U},\, \tilde{U}_x \in L^3(\mathbb{R}_T)$ (by $\tilde{U} \in W^{1,2}_{3/2}(\mathbb{R}_T)$ and the parabolic Sobolev inequality in the form of [28, chapter II, Lemma 3.3]) and the same argument as in step 2. Thus, by (4.32), the norm $\|\tilde{\Delta}\|_{W^{0,1}_2(\mathbb{R}_T)}$ tends to 0 in the limit $k \to \infty$ and we can conclude $U = \tilde{U}$ as desired. Finally, this, the representation $R_{xx} = U_{xx} + V_{xx} + Z_{xx}$ and $V_{xx},\, Z_{xx} \in L^{3/2}(\mathbb{R}_T)$ imply together that $R_{xx} \in L^{3/2}(\mathbb{R}_T)$. Moreover, due to $h \in L^2(\mathbb{R}_T, R_x\,\mathrm{d}t\,\mathrm{d}x)$, $R_x \in L^2(\mathbb{R}_T, R_x\,\mathrm{d}t\,\mathrm{d}x)$ and Hölder's inequality, it holds $h\,A(R)R_x \in L^{3/2}(\mathbb{R}_T)$, so that we also have $R_t = A'(R)R_x^2 + A(R)R_{xx} + h\,A(R)R_x \in L^{3/2}(\mathbb{R}_T)$. This finishes the proof. □

5. Proof of Proposition 2.2 (c)

The proof of Proposition 2.2 (C) is broken down into several steps: first, we show that the variational formula in the definition of $\tilde{I}$ (recall that $\tilde{I}(\gamma) \le I(\gamma)$, if $I(\gamma) < \infty$) can be rewritten as the one corresponding to a suitable one-dimensional reversible diffusion; then, we prove the existence of a sufficiently regular solution to the corresponding backward Cauchy problem, which enables us to employ Dirichlet form calculus to control the quantities of interest.

Proof of Proposition 2.2 (C). Step 1. We recall from the outline that $I(\gamma) < \infty$ implies $\tilde{I}(\gamma) \le I(\gamma) < \infty$, so that

$$\sup_{v \in \overline{\mathcal{S}}} \Big[(\gamma(T), v(T,\cdot)) - (\gamma(0), v(0,\cdot)) - \int_0^T (\gamma(t), \mathcal{R}^\gamma_t v + a\,(v_x)^2)\,\mathrm{d}t \Big] < \infty. \tag{5.1}$$

Moreover, by Proposition 2.2 (A), we have $R_x \in L^3(\mathbb{R}_T)$. Thus, by multiple applications of Hölder's inequality, we deduce that the functional in the supremum in (5.1) is continuous with respect to the norm

$$\|v\|_\infty + \|v_t\|_{L^{3/2}(\mathbb{R}_T)} + \|v_x\|_{L^{3/2}(\mathbb{R}_T)} + \|v_x\|_{L^3(\mathbb{R}_T)} + \|v_{xx}\|_{L^{3/2}(\mathbb{R}_T)}. \tag{5.2}$$

Therefore, may replace the supremum above by a supremum over functions $v \in C_b(\mathbb{R}_T)$ such that $v_t,\, v_{xx} \in L_{3/2}(\mathbb{R}_T)$ and $v_x \in L_{3/2}(\mathbb{R}_T) \cap L_3(\mathbb{R}_T)$. We denote the space of such functions by $\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$.

Now, let ψ be a function in $C^\infty(\mathbb{R})$ with $\lim_{|x|\to\infty} \frac{\psi(x)}{|x|} = 1$, $\|\psi'\|_\infty < \infty$ and such that $\alpha_0 := e^{-\psi(x)}\,\mathrm{d}x$ is a probability measure. We introduce next the parabolic

operator

$$\mathcal{R}_t^{\gamma,\psi} = \frac{\partial}{\partial t} + e^{\psi(x)}\frac{\partial}{\partial x}a(t,x)e^{-\psi(x)}\frac{\partial}{\partial x} = \mathcal{R}_t^{\gamma} + (a_x - a\psi' - b(R))\frac{\partial}{\partial x} \tag{5.3}$$

with $a = A(R)$ as before, and claim that the finiteness of the supremum in (5.1) over $v \in \tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$ implies that

$$\sup_{v\in\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)} \Big[(\gamma(T), v(T,\cdot)) - \log\int_{\mathbb{R}} e^{v(0,x)}\,\mathrm{d}\alpha_0 - \int_0^T (\gamma(t), \mathcal{R}_t^{\gamma,\psi}v + a\,(v_x)^2)\,\mathrm{d}t\Big] < \infty. \tag{5.4}$$

Indeed, the change from $\mathcal{R}_t^{\gamma}$ to $\mathcal{R}_t^{\gamma,\psi}$ does not make the supremum infinite, since by the Cauchy-Schwarz inequality we have the bound

$$\beta_3 := \int_{\mathbb{R}_T} (b(R) - a_x + a\psi')v_x\,R_x\,\mathrm{d}m \le C_2\sqrt{\beta_2},$$

with

$$C_2 := \big(\|b\|_\infty T^{1/2} + \|A'\|_\infty \|R_x\|^{3/2}_{L^3} + \|a\psi'\|_\infty T^{1/2}\big)/(\inf_{\mathbb{R}_T} a) \;\text{ and}$$
$$\beta_2 := \int_0^T (\gamma(t), a\,v_x^2)\,\mathrm{d}t,$$

(note that the function a is bounded away from 0), we can scale the test function v in the two suprema by arbitrary constants $\lambda > 0$ and $\tilde{\lambda} > 0$, and the implication

$$\begin{aligned} &\forall\,\lambda > 0:\; \lambda\beta_1 \le C_1 + \lambda^2\beta_2,\; \beta_3 \le C_2\sqrt{\beta_2} \\ \Rightarrow\;\; &\forall\,\tilde{\lambda} > 0:\; \tilde{\lambda}\beta_1 + \tilde{\lambda}\beta_3 \le 2C_1 + \frac{1}{2}C_2^2 + \tilde{\lambda}^2\beta_2 \end{aligned}$$

holds for all $\beta_1 \in \mathbb{R}$, $\beta_2 > 0$, $\beta_3 \in \mathbb{R}$, $C_1 > 0$ and $C_2 > 0$. Therefore, choosing

$$\beta_1 = (\gamma(T), v(T,\cdot)) - (\gamma(0), v(0,\cdot)) - \int_0^T (\gamma(t), \mathcal{R}_t^{\gamma} v)\,\mathrm{d}t,$$

$C_1 = \tilde{I}(\gamma)$ and $\lambda = \frac{\tilde{\lambda}}{2}$ in the latter implication, one justifies the change from $\mathcal{R}_t^{\gamma}$ to $\mathcal{R}_t^{\gamma,\psi}$. Moreover, the relative entropy

$$H(\rho_0|\alpha_0) = \int_{\mathbb{R}} \log\frac{\mathrm{d}\rho_0}{\mathrm{d}\alpha_0}\,\mathrm{d}\rho_0 = \int_{\mathbb{R}} \log\frac{\theta}{e^{-\psi}}\,\mathrm{d}\rho_0 \le \int_{\mathbb{R}} |\psi|\,\mathrm{d}\rho_0 + \int_{\mathbb{R}} \log\theta\,\mathrm{d}\rho_0 \tag{5.5}$$

is finite due to $\int_{\mathbb{R}} |x|\,\mathrm{d}\rho_0 < \infty$ and $\theta \in L^3(\mathbb{R})$ (see Assumption 2). Therefore, the change in the term corresponding to the initial condition can be justified by

$$\sup_{v\in\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)} \Big[\int_{\mathbb{R}} v(0,x)\rho_0(\mathrm{d}x) - \log\int_{\mathbb{R}} e^{v(0,x)}\,\mathrm{d}\alpha_0\Big] \le H(\rho_0|\alpha_0) < \infty.$$

Here, the first inequality follows from [10, Lemma 6.2.13].

Next, for each function $v \in \tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$, we introduce the function $u = e^v$, let $\mathcal{E}\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$ be the space of such functions and rewrite the supremum in (5.4) as

$$\sup_{u\in\mathcal{E}\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)} \Big[(\gamma(T), \log u(T,\cdot)) - \log \int_{\mathbb{R}} u(0,x)\,\mathrm{d}\alpha_0 - \int_0^T \Big(\gamma(t), \frac{\mathcal{R}^{\gamma,\psi}_t u}{u}\Big)\,\mathrm{d}t\Big]. \tag{5.6}$$

Step 2. We claim next that, for every function $f \in \bar{\mathcal{S}}$, there exists a function $u \in \mathcal{E}\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$ such that

$$\mathcal{R}^{\gamma,\psi}_t u - f\,u = 0, \quad u(T,\cdot) = 1. \tag{5.7}$$

Here, we can understand the PDE (5.7) as an equation in $L^{3/2}(\mathbb{R}_T)$ (since by the parabolic Sobolev inequality $u_x \in L^3(\mathbb{R}_T)$ for all $u \in \mathcal{E}\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$, so that $(a_x - a\psi')u_x \in L^{3/2}(\mathbb{R}_T)$ by the Cauchy-Schwarz inequality, the boundedness of the function A' and Proposition 2.2 (A)). To prove the claim, it clearly suffices to show that there exists a function w solving

$$\mathcal{R}^{\gamma,\psi}_t w - f\,w = f, \quad w(T,\cdot) = 0 \tag{5.8}$$

such that $w+1$ is an element of $\mathcal{E}\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$. To this end, we first rewrite the latter PDE as

$$w_t + (aw_x)_x - \psi'\,a\,w_x - f\,w = f, \quad w(T,\cdot) = 0, \tag{5.9}$$

and employ [26, Theorem 6.2] together with the method of continuity to find a generalized solution w of (5.9) in the space $W^{0,1}_6(\mathbb{R}_T) \cap W^{0,1}_2(\mathbb{R}_T)$. Indeed, the norm bound in [26, inequality (6.3)] extends to a norm bound for functions in $\mathcal{H}^1_6 \cap \mathcal{H}^1_2$ with respect to the norms $\|\cdot\|_{\mathbb{H}^{-1}_6} + \|\cdot\|_{\mathbb{H}^{-1}_2}$ and $\|\cdot\|_{\mathbb{H}^1_6} + \|\cdot\|_{\mathbb{H}^1_2}$ defined in [26] (one only needs to add the norm bounds [26, inequality (6.3)] for $p=6$ and $p=2$). Applying the method of continuity (see [29, section III.2]) relying on such a refined norm estimate to interpolate between the PDE (5.9) and the corresponding PDE with a smooth coefficient a, we find a solution of (5.9) in $\mathcal{H}^1_6 \cap \mathcal{H}^1_2$, which in particular belongs to $W^{0,1}_6(\mathbb{R}_T) \cap W^{0,1}_2(\mathbb{R}_T)$. Moreover, refining the norm bounds in [26, Theorem 4.1] to an estimate on the norms in $L^2(\mathbb{R}_T) \cap L^{3/2}(\mathbb{R}_T)$ and applying the method of continuity in a similar fashion, we can find a generalized solution $\hat{w} \in W^{1,2}_2(\mathbb{R}_T) \cap W^{1,2}_{3/2}(\mathbb{R}_T)$ of the equation

$$\hat{w}_t + a\,\hat{w}_{xx} - a\,\psi'\hat{w}_x - f\,\hat{w} = -a_x\,w_x + f, \quad \hat{w}(T,\cdot) = 0, \tag{5.10}$$

since Hölder's inequality and $a_x \in L^3(\mathbb{R}_T) \cap L^2(\mathbb{R}_T)$ (due to Proposition 2.2 (A) and its proof, as well as the boundedness of the function A') imply $a_x\,w_x \in L^2(\mathbb{R}_T) \cap L^{3/2}(\mathbb{R}_T)$. To show that $w = \hat{w}$, we let ϕ_k, $k \in \mathbb{N}$ be a truncation sequence as in the proof of Proposition 2.2 (B), fix a $k \in \mathbb{N}$ and set $\Delta := \phi_k(\hat{w} - w)$. Then, $\Delta \in W^{0,1}_2(\mathbb{R}_T)$ is a generalized solution of

$$\Delta_t + (a\Delta_x)_x - \psi'\,a\,\Delta_x - f\,\Delta = \psi_k, \quad \Delta(T,\cdot) = 0, \tag{5.11}$$

where

$$\psi_k = \phi_k''\, a(\hat{w} - w) + 2\,\phi_k'\, a(\hat{w} - w)_x - \phi_k'\, \psi'\, a(\hat{w} - w). \tag{5.12}$$

Recalling the notation $S_n = [0,T]\times[n,n+1]$, $n \in \mathbb{Z}$, we deduce from $\max(|\phi_k'|, |\phi_k''|) \leq 2 \cdot \mathbf{1}_{S_k \cup S_{-k-1}}$, the boundedness of a and $(\hat{w} - w), (\hat{w} - w)_x \in L^2(\mathbb{R}_T)$ that

$$\lim_{k\to\infty} \|\psi_k\|_{L^2(\mathbb{R}_T)} = 0. \tag{5.13}$$

Hence, writing $\Delta = \phi_k\, \hat{w} - \phi_k\, w$, following the paragraph after the statement of [28, chapter III, Theorem 3.3] and applying the energy inequality of [28, chapter III, Theorem 2.1], we conclude that $\|\Delta\|_{L^2(\mathbb{R}_T)} \to 0$ as $k \to \infty$. Therefore, $w = \hat{w}$ Lebesgue almost everywhere. All in all, we have found a solution u to (5.7) such that $u - 1$ is an element of $W_2^{1,2}(\mathbb{R}_T) \cap W_{3/2}^{1,2}(\mathbb{R}_T)$. We also note at this point that the parabolic Sobolev inequality in the form of [28, chapter II, Lemma 3.3] with $p = 6$, $q = 2$ applied to $\hat{w}$ gives $u_x = w_x = \hat{w}_x \in L^6(\mathbb{R}_T)$.

It remains to show that $u \in C_b(\mathbb{R}_T)$ and that u is bounded away from 0 on $\mathbb{R}_T$. To this end, we first apply [28, chapter III, Theorem 5.2] to find a generalized solution $\tilde{w}$ of (5.9) in the subspace of $W_2^{0,1}(\mathbb{R}_T)$, whose elements satisfy

$$\operatorname{ess\,sup}_{t\in[0,T]} \int_{\mathbb{R}} \tilde{w}(t,x)^2\, \mathrm{d}x + \int_{\mathbb{R}_T} \tilde{w}_x^2\, \mathrm{d}m < \infty. \tag{5.14}$$

Next, we apply [1, Theorem 10 (vi)] with the constant $\gamma > 0$ there being arbitrarily small to conclude that $\tilde{u} := \tilde{w} + 1$ has to be the unique generalized solution of (5.9) in the sense of [1] and is given by

$$\tilde{u}(t,x) = \int_{\mathbb{R}} \Gamma(t,x;T,y)\, \mathrm{d}y \tag{5.15}$$

with Γ being the weak fundamental solution of the equation (5.7) defined as in [1, section 6]. Now, [1, Theorem C] implies that $\tilde{u}$ is locally Hölder continuous in (t,x), hence, also continuous in (t,x) on $[0,T)\times\mathbb{R}$. Putting this together with [1, Theorem 10 (vi)], we conclude that $\tilde{u}$ is continuous on the whole of $\mathbb{R}_T$. Finally, we use the heat kernel estimates on Γ of [1, Theorem 7] to conclude that $\tilde{u}$ has to be bounded between two positive constants. Therefore, all we need to show now is that $\tilde{u} = u$, or equivalently $\tilde{w} = w$. To this end, we let ϕ_k, $k \in \mathbb{N}$ be a truncation sequence as above, fix a $k \in \mathbb{N}$ and set $\tilde{\Delta} := \phi_k(\tilde{w} - w)$. Then, $\tilde{\Delta} \in W_2^{0,1}(\mathbb{R}_T)$ is a generalized solution of

$$\tilde{\Delta}_t + (a\tilde{\Delta}_x)_x - \psi'\, a\, \tilde{\Delta}_x - f\, \tilde{\Delta} = \tilde{\psi}_k, \quad \tilde{\Delta}(T,\cdot) = 0, \tag{5.16}$$

where

$$\tilde{\psi}_k = \phi_k''\, a(\tilde{w} - w) + 2\,\phi_k'\, a(\tilde{w} - w)_x + \phi_k'\, a_x\, (\tilde{w} - w) - \phi_k'\, \psi'\, a(\tilde{w} - w). \tag{5.17}$$

In addition, from $w, \tilde{w} \in W_2^{0,1}(\mathbb{R}_T)$ we deduce

$$\lim_{n\to\pm\infty} \|\tilde{w} - w\|_{L^2(S_n)} = \lim_{n\to\pm\infty} \|\tilde{w}_x - w_x\|_{L^2(S_n)} = 0. \tag{5.18}$$

Putting this together with $\max(|\phi_k'|,|\phi_k''|)\le 2\cdot\mathbf{1}_{S_k\cup S_{-k-1}}$, the uniform boundedness of the norms $\|\tilde{w}-w\|_{L^6(S_n)}$, $n\in\mathbb{Z}$, the inclusion $a_x\in L^3(\mathbb{R}_T)$ (see Proposition 2.2 (A)) and Hölder's inequality, we can conclude

$$\lim_{k\to\infty}\|\tilde{\psi}_k\|_{L^2(\mathbb{R}_T)}=0. \tag{5.19}$$

Hence, writing $\tilde{\Delta}=\phi_k\,\tilde{w}-\phi_k\,w$, following the paragraph after the statement of [28, chapter III, Theorem 3.3] and applying the energy inequality of [28, chapter III, Theorem 2.1], we deduce that $\|\tilde{\Delta}\|_{L^2(\mathbb{R}_T)}\to 0$ as $k\to\infty$. Hence, $w=\tilde{w}$ Lebesgue almost everywhere, as desired.

Step 3. In the previous steps we have shown that

$$\sup_{f\in\bar{\mathcal{S}}}\Big[-\log\int_{\mathbb{R}}u(0,x)\,\mathrm{d}\alpha_0-\int_0^T(\gamma(t),f)\,\mathrm{d}t\Big]<\infty, \tag{5.20}$$

where u is the function in $\mathcal{E}\tilde{W}^{1,2}_{3/2}(\mathbb{R}_T)$ corresponding to f via (5.7). We next take $f=g_x+C\,g^2$ with $g\in\overline{\mathcal{S}}$ and will bound the corresponding term $\int_{\mathbb{R}}u(0,x)\,\mathrm{d}\alpha_0$ from above. Approximating the solution u of the PDE (5.7) by functions in $\overline{\mathcal{S}}$ converging to u in $\tilde{W}^{1,2}_{3/2}$, we deduce that, for all $0\le t_1<t_2\le T$:

$$\begin{aligned}
&\|u(t_2,\cdot)\|^2_{L^2(\alpha_0)}-\|u(t_1,\cdot)\|^2_{L^2(\alpha_0)}=2\int_{t_1}^{t_2}\int_{\mathbb{R}}u_t\,u\,\mathrm{d}\alpha_0\,\mathrm{d}t\\
=\;&-2\int_{t_1}^{t_2}\int_{\mathbb{R}}(\mathcal{A}_t\,u-f\,u)u\,\mathrm{d}\alpha_0\,\mathrm{d}t\ge -2\int_{t_1}^{t_2}\lambda(t)\,\|u(t,\cdot)\|^2_{L^2(\alpha_0)}\,\mathrm{d}t,
\end{aligned}$$

where

$$\begin{aligned}
\lambda(t)&=\sup_{h_1\in C_c^\infty(\mathbb{R}):\,h_1\ge 0,\,\|h_1\|_{L^2(\alpha_0)}=1}\int_{\mathbb{R}}(\mathcal{A}_t\,h_1-f\,h_1)\,h_1\,\mathrm{d}\alpha_0\\
&=\sup_{h_1\in C_c^\infty(\mathbb{R}):\,h_1\ge 0,\,\|h_1\|_{L^2(\alpha_0)}=1}-\int_{\mathbb{R}}a\,(h_1)_x^2\,\mathrm{d}\alpha_0-\int_{\mathbb{R}}f\,h_1^2\,\mathrm{d}\alpha_0\\
&=\sup_{h_2\in C_c^\infty(\mathbb{R}):\,h_2\ge 0,\,\|h_2\|_{L^1(\alpha_0)}=1}\Big[-\frac14\int_{\mathbb{R}}\frac{(h_2')^2}{h_2}\,a\,\mathrm{d}\alpha_0-\int_{\mathbb{R}}f\,h_2\,\mathrm{d}\alpha_0\Big]
\end{aligned}$$

and we took $h_1=\sqrt{h_2}$. Moreover, using integration by parts followed by the Cauchy-Schwarz inequality, we conclude

$$\Big|\int_{\mathbb{R}}g_x\,h_2\,\mathrm{d}\alpha_0\Big|=\Big|\int_{\mathbb{R}}g(h_2'-h_2\psi')\,\mathrm{d}\alpha_0\Big|\le\sqrt{\upsilon_1\upsilon_2}+\|\psi'\|_\infty\sqrt{\upsilon_2}, \tag{5.21}$$

where we have set $\upsilon_1=\int_{\mathbb{R}}\frac{(h_2')^2}{h_2}\,\mathrm{d}\alpha_0$, $\upsilon_2=\int_{\mathbb{R}}g^2\,h_2\,\mathrm{d}\alpha_0$. Hence, we end up with

$$\lambda(t)\le\sup_{\upsilon_1,\upsilon_2\ge 0}\Big[-\frac{(\inf_{\mathbb{R}_T}a)\upsilon_1}{4}+\sqrt{\upsilon_1\upsilon_2}+\|\psi'\|_\infty\sqrt{\upsilon_2}-C\upsilon_2\Big],\quad t\in[0,T]. \tag{5.22}$$

We now fix $C>0$ large enough, so that the right side of (5.22) is some $\tilde{C}<\infty$. Next, we apply Gronwall's Lemma to the function $t\mapsto\|u(t,\cdot)\|^2_{L^2(\alpha_0)}$ to conclude that $\|u(0,\cdot)\|^2_{L^2(\alpha_0)}\le e^{2\tilde{C}T}$. Thus, we have just shown that

$$\log\int_{\mathbb{R}}u(0,x)\,\mathrm{d}\alpha_0\le\log\|u(0,\cdot)\|_{L^2(\alpha_0)}\le\tilde{C}T. \tag{5.23}$$

Putting this together with (5.20) and integrating by parts in space, we end up with

$$\infty>\sup_{g\in\bar{\mathcal{S}}}\int_0^T(\gamma(t),-g_x-C\,g^2)\,\mathrm{d}t=\sup_{g\in\bar{\mathcal{S}}}\int_{\mathbb{R}_T}(g\,R_{xx}-C\,g^2\,R_x)\,\mathrm{d}m. \tag{5.24}$$

Finally, we consider the Hilbert space $\tilde{\mathbb{H}}$ given by the closure of $\bar{\mathcal{S}}$ with respect to the norm $\|g\|_{\tilde{\mathbb{H}}}=\int_{\mathbb{R}_T}g^2\,R_x\,\mathrm{d}m$ on $\bar{\mathcal{S}}$. We have shown that $g\mapsto\int_{\mathbb{R}_T}g\,R_{xx}\,\mathrm{d}m$ is a bounded linear functional on $\bar{\mathcal{S}}$ with respect to $\|\cdot\|_{\tilde{\mathbb{H}}}$. Thus, it extends to a bounded linear functional on $\tilde{\mathbb{H}}$ and, by the Riesz Representation Theorem, can be represented as $g\mapsto(g,\tilde{h})_{\tilde{\mathbb{H}}}$ for a suitable $\tilde{h}\in\tilde{\mathbb{H}}$. Moreover, the identity $\int_{\mathbb{R}_T}g\,R_{xx}\,\mathrm{d}m=\int_{\mathbb{R}_T}g\,\tilde{h}\,R_x\,\mathrm{d}m$ for all $g\in\bar{\mathcal{S}}$ shows that $\tilde{h}$ can be chosen as $\frac{R_{xx}}{R_x}$ with the convention $\frac{0}{0}=0$. Therefore, the supremum on the right side of (5.24) can be written as

$$\sup_{g\in\tilde{\mathbb{H}}}\big((g,\tilde{h})_{\tilde{\mathbb{H}}}-C(g,g)_{\tilde{\mathbb{H}}}\big)=\frac{1}{4C}(\tilde{h},\tilde{h})_{\tilde{\mathbb{H}}}=\int_{\mathbb{R}_T}\frac{R_{xx}^2}{R_x}\,\mathrm{d}m, \tag{5.25}$$

and we conclude $\int_{\mathbb{R}_T}\frac{R_{xx}^2}{R_x}\,\mathrm{d}m<\infty$, as desired.

Step 4. We now turn to the proof of $\int_{\mathbb{R}_T}\frac{R_t^2}{R_x}\,\mathrm{d}m<\infty$. As before, it suffices to show

$$\sup_{g\in\bar{\mathcal{S}}}\int_{\mathbb{R}_T}(g\,R_t-C\,g^2\,R_x)\,\mathrm{d}m<\infty. \tag{5.26}$$

Moreover, the PDE $R_t=(A(R)R_x)_x+h\,A(R)\,R_x$ allows to rewrite the latter supremum as

$$\sup_{g\in\bar{\mathcal{S}}}\int_{\mathbb{R}_T}\big(A'(R)R_x^2+A(R)R_{xx}+h\,A(R)\,R_x\big)g\,\mathrm{d}m-\int_{\mathbb{R}_T}C\,g^2\,R_x\,\mathrm{d}m. \tag{5.27}$$

Moreover, optimizing pointwise we see that the supremum cannot exceed $C\,(\hat{h},\hat{h})_{\tilde{\mathbb{H}}}$, where

$$\hat{h}=\frac{A'(R)R_x^2+A(R)R_{xx}+h\,A(R)\,R_x}{2\,C\,R_x}\in\tilde{\mathbb{H}} \tag{5.28}$$

with our usual convention $\frac{0}{0}=0$. In particular, the supremum of interest is finite. □

Remark 5.1. A crucial step in the proof of Proposition 2.2 (C) consists of showing that the solution of (5.7) in $W_2^{1,2}(\mathbb{R}_T)\cap C(\mathbb{R}_T)$ is bounded and bounded away from 0. Instead of relying on the results of [1], this can be also seen from [27, Corollary 4.6]. Indeed, for each $n\in\mathbb{N}$, let u_n be the solution of the PDE in (5.7) on $[0,T]\times[-n,n]$, which belong to $W_2^{1,2}([0,T]\times[-n,n])\cap C([0,T]\times[-n,n])$ and satisfies the boundary conditions $u_n(\cdot,\pm n)=1$, $u_n(T,\cdot)=1$. Such solutions can be constructed as in step

2 in the proof of Proposition 2.2 (C) by first employing the results of [1] to find a solution in $C([0,T]\times[-n,n])$, then by using the results in [26] to find a solution in $W_2^{1,2}([0,T]\times[-n,n])$ and finally by appealing to [28, chapter III, Theorem 3.3] to conclude that the two solutions are the same. At this point, applying [27, Corollary 4.6] to the functions $v_n := e^{-\|f\|_\infty(T-t)} - u_n$, $n\in\mathbb{N}$ and $w_n := u_n - e^{\|f\|_\infty(T-t)}$, $n\in\mathbb{N}$ satisfying

$$(v_n)_t + (a(v_n)_x)_x - \psi'\,a\,(v_n)_x + f\,v_n = -(\|f\|_\infty + f)\,e^{-\|f\|_\infty(T-t)}, \tag{5.29}$$

$$(w_n)_t + (a(w_n)_x)_x - \psi'\,a\,(w_n)_x - f\,w_n = -(\|f\|_\infty - f)\,e^{\|f\|_\infty(T-t)} \tag{5.30}$$

almost everywhere, one obtains the inequalities

$$e^{-\|f\|_\infty(T-t)} \le u_n(t,x) \le e^{\|f\|_\infty(T-t)} \quad \text{for a.e. } (t,x)\in[0,T]\times[-n,n]. \tag{5.31}$$

We note here that one may assume that the sign of f is as needed to apply [27, Corollary 4.6], since otherwise we can replace v_n and w_n by $e^{c_1(T-t)}v_n$ and $e^{c_2(T-t)}w_n$, respectively, with suitable constants c_1, $c_2\in\mathbb{R}$. Moreover, the energy inequality [28, chapter III, Theorem 2.1] implies that the $W_2^{0,1}(\mathbb{R}_T)$ norms of the functions u_n-1, $n\in\mathbb{N}$ are uniformly bounded. Therefore, by the Banach-Alaoglu Theorem there exists a subsequence of the sequence u_n-1, $n\in\mathbb{N}$, which converges in weak-$W_2^{0,1}(\mathbb{R}_T)$ sense to a limit $u-1\in W_2^{0,1}(\mathbb{R}_T)$. It follows that u is a generalized solution of (5.7) on $\mathbb{R}_T$ with boundary condition $u(T,\cdot)=1$, which satisfies the inequalitites

$$e^{-\|f\|_\infty(T-t)} \le u(t,x) \le e^{\|f\|_\infty(T-t)} \quad \text{for a.e. } (t,x)\in[0,T]\times\mathbb{R}. \tag{5.32}$$

Finally, by [28, chapter III, Theorem 3.3] and the same truncation argument as in step 2 of the proof of Proposition 2.2 (C), it must coincide with the solution of (5.7) in $W_2^{1,2}([0,T]\times[-n,n])\cap C([0,T]\times[-n,n])$ found there.

6. Proof of Proposition 2.3

Throughout this section, we consider paths $\gamma\in C([0,T],M_1(\mathbb{R}))$ with $J(\gamma)<\infty$ and, with a small abuse of notation, write $J(R)$ for $J(\gamma)$. Before proving Proposition 2.3, we will show the following preliminary result.

Proposition 6.1. *Let $\gamma\in C([0,T],M_1(\mathbb{R}))$ be such that $J(\gamma)<\infty$. Then there exists a family of functions γ^ϵ, $\epsilon\in(0,1)$ in $C^{1,1}(\mathbb{R}_T)$ such that $\gamma^\epsilon(0,\cdot)=\frac{\mathrm{d}\rho_0}{\mathrm{d}x}$ and $\gamma^\epsilon(\cdot,x)\,\mathrm{d}x\in C([0,T],M_1(\mathbb{R}))$ for all $\epsilon\in(0,1)$, $\gamma^\epsilon(\cdot,x)\,\mathrm{d}x\to\gamma$ in $C([0,T],M_1(\mathbb{R}))$ as $\epsilon\downarrow 0$, and*

$$\lim_{\epsilon\downarrow 0} J(\gamma^\epsilon) = J(\gamma). \tag{6.1}$$

The proof of Proposition 6.1 is based on the following lemma. It shows the *convexity* of the functionals, which comprise the rate function J.

Lemma 6.2. *The functionals*

$$\tilde{R}\mapsto\int_{\mathbb{R}_T}\frac{\tilde{R}_t^2}{\tilde{R}_x}\,\mathrm{d}m,\quad \tilde{R}\mapsto\int_{\mathbb{R}_T}\frac{\tilde{R}_{xx}^2}{\tilde{R}_x}\,\mathrm{d}m,\quad \tilde{R}\mapsto\int_{\mathbb{R}_T}\tilde{R}_x^3\,\mathrm{d}m \tag{6.2}$$

on the set

$$\tilde{\mathcal{R}} := \Big\{ \tilde{R} = F_{\gamma(\cdot)}(\cdot) :\ \gamma \in C([0,T], M_1(\mathbb{R})), \tilde{R}_x \in L^3(\mathbb{R}_T), \tilde{R}_t, \tilde{R}_{xx} \in L^{3/2}(\mathbb{R}_T), \frac{\tilde{R}_t^2}{\tilde{R}_x}, \frac{\tilde{R}_{xx}^2}{\tilde{R}_x} \in L^1(\mathbb{R}_T) \Big\}$$

are all convex.

Proof. Arguing as at the end of step 3 in the proof of Proposition 2.2 (C), we conclude that on $\tilde{\mathcal{R}}$

$$\int_{\mathbb{R}_T} \frac{\tilde{R}_t^2}{\tilde{R}_x}\,\mathrm{d}m = \sup_{g\in\bar{\bar{S}}} \int_{\mathbb{R}_T} \big(2\,g\,\tilde{R}_t - g^2\tilde{R}_x\big)\,\mathrm{d}m,$$

$$\int_{\mathbb{R}_T} \frac{\tilde{R}_{xx}^2}{\tilde{R}_x}\,\mathrm{d}m = \sup_{g\in\bar{\bar{S}}} \int_{\mathbb{R}_T} \big(2\,g\,\tilde{R}_{xx} - g^2\tilde{R}_x\big)\,\mathrm{d}m.$$

In other words, each of the first two functionals is given by a supremum of linear functionals and is therefore convex. Moreover, the convexity of the third functional follows immediately from the convexity of the function $x \mapsto x^3$ on $[0,\infty)$. □

In the proof of Proposition 6.1 we will use the following notation. We fix functions $k \in C_0^\infty(\mathbb{R})$ and $\phi \in C_0^\infty([0,1])$ taking only positive values on $\mathbb{R}$ and $(0,1)$, respectively, such that $\int_{\mathbb{R}} k(x)\,\mathrm{d}x = \int_{[0,1]} \phi(x)\,\mathrm{d}x = 1$. Then, for every $\epsilon \in (0,1)$, we first define the function

$$S^\epsilon(t,x) = \begin{cases} R(0,x), & t\in[0,\epsilon] \\ \frac{2\epsilon-t}{\epsilon} R(0,x) + \frac{t-\epsilon}{\epsilon}\int_{\mathbb{R}} R(0,x+y)\,\delta^{-1}\,k(y/\delta)\,\mathrm{d}y, & t\in(\epsilon,2\epsilon] \\ \int_{\mathbb{R}} R(t-2\epsilon,x+y)\,\delta^{-1}\,k(y/\delta)\,\mathrm{d}y, & t\in(2\epsilon,T+\epsilon] \end{cases} \tag{6.3}$$

for some $\delta = \delta(\epsilon) \in (0,1)$ to be determined later. Subsequently, we introduce the functions

$$R^\epsilon(t,x) = \int_0^\epsilon S^\epsilon(t+s,x)\,\epsilon^{-1}\,\phi(s/\epsilon)\,\mathrm{d}s, \quad \epsilon\in(0,1) \tag{6.4}$$

on $\mathbb{R}_T$ and let $\gamma^\epsilon = R_x^\epsilon$, $\epsilon \in (0,1)$. Then, the functions γ^ϵ, $\epsilon \in (0,1)$ belong to the space $C^{1,1}(\mathbb{R}_T)$, by Fubini's Theorem $\gamma^\epsilon(.,x)\,\mathrm{d}x$ is an element of $C([0,T], M_1(\mathbb{R}))$, and $\gamma^\epsilon(.,x)\,\mathrm{d}x \to \gamma$ in $C([0,T], M_1(\mathbb{R}))$ as ϵ goes to 0, provided that $\delta(\epsilon) \to 0$ in the same limit. Thus, to prove Proposition 6.1, it suffices to show (6.1) along a subsequence.

Proof of Proposition 6.1. Step 1. We first claim that

$$\limsup_{\epsilon\downarrow 0} \int_{\mathbb{R}_T} \frac{(R_t^\epsilon)^2}{R_x^\epsilon}\,\mathrm{d}m \le \int_{\mathbb{R}_T} \frac{R_t^2}{R_x}\,\mathrm{d}m, \tag{6.5}$$

$$\limsup_{\epsilon\downarrow 0} \int_{\mathbb{R}_T} \frac{(R_{xx}^\epsilon)^2}{R_x^\epsilon}\,\mathrm{d}m \le \int_{\mathbb{R}_T} \frac{R_{xx}^2}{R_x}\,\mathrm{d}m, \tag{6.6}$$

$$\limsup_{\epsilon\downarrow 0} \int_{\mathbb{R}_T} (R_x^\epsilon)^3\,\mathrm{d}m \le \int_{\mathbb{R}_T} R_x^3\,\mathrm{d}m. \tag{6.7}$$

Using the definitions of R^ϵ, S^ϵ and applying Lemma 6.2 twice, we obtain

$$\int_{\mathbb{R}_T} \frac{(R^\epsilon_t)^2}{R^\epsilon_x}\,\mathrm{d}m \le \int_0^\epsilon \epsilon^{-1}\phi(s/\epsilon)\int_{[\epsilon-s,2\epsilon-s]\times\mathbb{R}} \frac{(R_0-\bar{R}^\epsilon_0)^2}{\epsilon\big((2\epsilon-t)R_0'+(t-\epsilon)(\bar{R}^\epsilon_0)'\big)}\,\mathrm{d}m\,\mathrm{d}s + \int_{\mathbb{R}_T}\frac{R_t^2}{R_x}\,\mathrm{d}m.$$

Here, we have defined $R_0 := R(0,\cdot)$, $\bar{R}^\epsilon_0 := \int_{\mathbb{R}} R(0,\cdot+y)\,\delta^{-1}\,k(y/\delta)\,\mathrm{d}y$. Writing $R_0-\bar{R}^\epsilon_0$ as $(1-\bar{R}^\epsilon_0)-(1-R_0)$ on $(0,\infty)$, we conclude from Assumption 2 that by choosing $\delta=\delta(\epsilon)\in(0,1)$ small enough, we can make the latter upper bound arbitrarily close to the right side of (6.5), so that (6.5) readily follows. By the same arguments we obtain

$$\begin{aligned}\int_{\mathbb{R}_T} \frac{(R^\epsilon_{xx})^2}{R^\epsilon_x}\,\mathrm{d}m &\le \int_0^\epsilon \epsilon^{-1}\phi(s/\epsilon)\int_{[0,\epsilon-s]\times\mathbb{R}} \frac{(R_0'')^2}{R_0'}\,\mathrm{d}m\,\mathrm{d}s \\ &\quad + \int_0^\epsilon \epsilon^{-1}\phi(s/\epsilon)\int_{[\epsilon-s,2\epsilon-s]\times\mathbb{R}} \frac{\big(\frac{2\epsilon-t}{\epsilon}R_0''+\frac{t-\epsilon}{\epsilon}(\bar{R}^\epsilon_0)''\big)^2}{\frac{2\epsilon-t}{\epsilon}R_0'+\frac{t-\epsilon}{\epsilon}(\bar{R}^\epsilon_0)'}\,\mathrm{d}m\,\mathrm{d}s + \int_{\mathbb{R}_T}\frac{R_{xx}^2}{R_x}\,\mathrm{d}m, \\ \int_{\mathbb{R}_T}(R^\epsilon_x)^3\,\mathrm{d}m &\le \int_0^\epsilon \epsilon^{-1}\phi(s/\epsilon)\int_{[0,\epsilon-s]\times\mathbb{R}}(R_0')^3\,\mathrm{d}m\,\mathrm{d}s \\ &\quad + \int_0^\epsilon \epsilon^{-1}\phi(s/\epsilon)\int_{[\epsilon-s,2\epsilon-s]\times\mathbb{R}}\Big(\frac{2\epsilon-t}{\epsilon}R_0'+\frac{t-\epsilon}{\epsilon}(\bar{R}^\epsilon_0)'\Big)^3\,\mathrm{d}m\,\mathrm{d}s + \int_{\mathbb{R}_T}R_x^3\,\mathrm{d}m.\end{aligned}$$

Again, we conclude from Assumption 2 that by choosing $\delta=\delta(\epsilon)\in(0,1)$ small enough, we can make the two upper bounds arbitrarily close to the right sides of (6.6), (6.7), respectively, so that (6.6) and (6.7) must hold.

Step 2. In this step we will show that as $\epsilon\downarrow 0$:

$$\int_{\mathbb{R}_T}\Big|\frac{R^\epsilon_t}{(R^\epsilon_x)^{1/2}}-\frac{R_t}{(R_x)^{1/2}}\Big|^2\,\mathrm{d}m\to 0, \tag{6.8}$$

$$\int_{\mathbb{R}_T}\Big|\frac{R^\epsilon_{xx}}{(R^\epsilon_x)^{1/2}}-\frac{R_{xx}}{(R_x)^{1/2}}\Big|^2\,\mathrm{d}m\to 0, \tag{6.9}$$

$$\int_{\mathbb{R}_T}\big|(R^\epsilon_x)^{3/2}-(R_x)^{3/2}\big|^2\,\mathrm{d}m\to 0. \tag{6.10}$$

To this end, we note first that the triangle inequality in the form

$$\big|(R^\epsilon_x)^{1/2}-(R_x)^{1/2}\big|^2 \le \big((R^\epsilon_x)^{1/2}+(R_x)^{1/2}\big)\cdot\big|(R^\epsilon_x)^{1/2}-(R_x)^{1/2}\big| = |R^\epsilon_x-R_x| \tag{6.11}$$

and the convergence $R^\epsilon_x\to R_x$ in $L^2(\mathbb{R}_T)$ as $\epsilon\downarrow 0$ (since $R_x\in L^2(\mathbb{R}_T)$ due to the proof of Proposition 2.2 (A)) imply

$$\int_{\mathbb{R}_T}\big|(R^\epsilon_x)^{1/2}-(R_x)^{1/2}\big|^2\,\mathrm{d}m\to 0 \tag{6.12}$$

as $\epsilon\downarrow 0$. Next, we introduce the notation $c^\epsilon=\frac{R^\epsilon_t}{(R^\epsilon_x)^{1/2}}$, and conclude from (6.5) and the Banach-Alaoglu Theorem that $c^\epsilon\to c^*$ weakly in $L^2(\mathbb{R}_T)$ along a subsequence for some $c^*\in L^2(\mathbb{R}_T)$. Without loss of generality we may assume that the whole

sequence converges to c^*. Combining the two latter observations with the triangle inequality, we conclude that for every $\psi \in C_c(\mathbb{R}_T)$:

$$\int_{\mathbb{R}_T} (R^\epsilon_x)^{1/2} c^\epsilon \,\psi \,\mathrm{d}m \to \int_{\mathbb{R}_T} (R_x)^{1/2} c^* \,\psi\,\mathrm{d}m, \quad \epsilon \downarrow 0. \tag{6.13}$$

Moreover, from Hölder's inequality, $(R_x)^{1/2} \in L^6(\mathbb{R}_T)$ (due to Proposition 2.2 (A)) and $c^* \in L^2(\mathbb{R}_T)$ we deduce $(R_x)^{1/2} c^* \in L^{3/2}(\mathbb{R}_T)$. On the other hand, from $(R^\epsilon_x)^{1/2} c^\epsilon = R^\epsilon_t$ and $R^\epsilon_t \to R_t$, $\epsilon \downarrow 0$ (as a consequence of Proposition 2.2 (B)), we conclude that for every $\psi \in C_c(\mathbb{R}_T)$:

$$\int_{\mathbb{R}_T} (R^\epsilon_x)^{1/2} c^\epsilon \,\psi\,\mathrm{d}m \to \int_{\mathbb{R}_T} R_t \,\psi\,\mathrm{d}m, \quad \epsilon \downarrow 0. \tag{6.14}$$

Putting (6.13) and (6.14) together we conclude $c^* = \frac{R_t}{(R_x)^{1/2}}$ and, hence, $\frac{R^\epsilon_t}{(R^\epsilon_x)^{1/2}} \to \frac{R_t}{(R_x)^{1/2}}$ weakly in $L^2(\mathbb{R}_T)$ as $\epsilon \downarrow 0$. This together with (6.5) yields (6.8). Moreover, the same line of proof shows (6.9).

To show (6.10), we start with the elementary inequality

$$|x_1^{3/2} - x_2^{3/2}| \le \frac{3}{2} |x_2 - x_1| \max(x_1, x_2)^{1/2} \tag{6.15}$$

for all x_1, x_2 in $[0, \infty)$. Combining this and Hölder's inequality we have

$$\begin{aligned}\int_{\mathbb{R}_T} \big|(R^\epsilon_x)^{3/2} - (R_x)^{3/2}\big|^2 \,\mathrm{d}m &\le \frac{9}{4} \int_{\mathbb{R}_T} |R^\epsilon_x - R_x|^2 \max(R^\epsilon_x, R_x) \mathrm{d}m \\ &\le \frac{9}{4} \Big(\int_{\mathbb{R}_T} |R^\epsilon_x - R_x|^3 \,\mathrm{d}m\Big)^{2/3} \Big(\int_{\mathbb{R}_T} \max(R^\epsilon_x, R_x)^3 \,\mathrm{d}m\Big)^{1/3}.\end{aligned}$$

The latter expression tends to 0 in the limit $\epsilon \downarrow 0$ due to $R^\epsilon_x \to R_x$, $\epsilon \downarrow 0$ in $L^3(\mathbb{R}_T)$ (as a consequence of Proposition 2.2 (A)) and (6.7). Thus, we have shown (6.10).

Step 3. To deduce (6.1) from (6.8), (6.9), (6.10) and (6.12), we recall first that $J(R) < \infty$ implies that R is continuous in (t, x) by the definition of the space $\mathcal{F}$ in the introduction. Hence, $R^\epsilon \to R$, $\epsilon \downarrow 0$ uniformly on compact sets. In addition, since $\gamma \in C([0, T], M_1(\mathbb{R}))$, the set $\{\gamma(t) : t \in [0, T]\}$ is compact, so by Prokhorov's Theorem the family $\gamma(t)$, $t \in [0, T]$ is uniformly tight. Thus, for every $\delta > 0$, there is a $K \in \mathbb{R}$ such that

$$\forall t \in [0, T],\, x \ge K:\ R(t, x) \ge 1 - \delta \quad \text{and} \quad \forall t \in [0, T],\, x \le -K:\ R(t, x) \le \delta. \tag{6.16}$$

The uniformity of the convergence $R^\epsilon \to R$, $\epsilon \downarrow 0$ on compact sets, the monotonicity of the functions R^ϵ, $\epsilon \in (0, 1)$ and R in x and (6.16) imply together that the functions R^ϵ, $\epsilon \in (0, 1)$ converge to the function R uniformly on $\mathbb{R}_T$ as $\epsilon \downarrow 0$. This and the continuity of the functions σ, A' and b on $[0, 1]$ (see Assumptions 1, 2) shows that the convergences

$$\sigma(R^\epsilon) \to \sigma(R), \quad A'(R^\epsilon) \to A'(R), \quad b(R^\epsilon) \to b(R) \tag{6.17}$$

as $\epsilon \downarrow 0$ are uniform on $\mathbb{R}_T$. Moreover, all functions involved in the latter three convergences are uniformly bounded on $\mathbb{R}_T$. Putting this together with the uniform positivity of σ, (6.8), (6.9), (6.10) and (6.12), we have the following convergences in $L^2(\mathbb{R}_T)$ in the limit $\epsilon \downarrow 0$:

$$\frac{1}{\sigma(R^\epsilon)}\,\frac{R^\epsilon_t}{(R^\epsilon_x)^{1/2}} \to \frac{1}{\sigma(R)}\,\frac{R_t}{(R_x)^{1/2}}, \tag{6.18}$$

$$\sigma(R^\epsilon)\frac{R^\epsilon_{xx}}{(R^\epsilon_x)^{1/2}} \to \sigma(R)\frac{R_{xx}}{(R_x)^{1/2}}, \tag{6.19}$$

$$\frac{A'(R^\epsilon)}{\sigma(R^\epsilon)}\,(R^\epsilon_x)^{3/2} \to \frac{A'(R)}{\sigma(R)}\,(R_x)^{3/2}, \tag{6.20}$$

$$\frac{b(R^\epsilon)}{\sigma(R^\epsilon)}(R^\epsilon_x)^{1/2} \to \frac{b(R)}{\sigma(R)}(R_x)^{1/2}. \tag{6.21}$$

Thus, as $\epsilon \downarrow 0$:

$$\begin{aligned} J(R^\epsilon) \;&=\; \frac{1}{2}\Big\|\frac{R^\epsilon_t-(A(R^\epsilon)R^\epsilon_x)_x}{\sigma(R^\epsilon)(R^\epsilon_x)^{1/2}}+\frac{b(R^\epsilon)}{\sigma(R^\epsilon)}(R^\epsilon_x)^{1/2}\Big\|^2_{L^2(\mathbb{R}_T)} \\ &\longrightarrow \frac{1}{2}\Big\|\Big[\frac{R_t-(A(R)R_x)_x}{\sigma(R)(R_x)^{1/2}}+\frac{b(R)}{\sigma(R)}(R_x)^{1/2}\Big]\,\mathbf{1}_{\{R_x>0\}}\Big\|^2_{L^2(\mathbb{R}_T)} = J(R). \end{aligned}$$

This finishes the proof of the proposition. $\square$

We are now ready to give the proof of Proposition 2.3.

Proof of Proposition 2.3. Fix a γ as in the statement of the proposition and the corresponding function R. In view of Proposition 6.1, we may assume that $R=\tilde{R}*\bar{k}^\epsilon$ for some $\tilde{R}$ with $J(\tilde{R})<\infty$ and some $\epsilon\in(0,1)$. Then, for each $M>0$, there is a constant $r^M\in(0,1)$ small enough, for which there exists a function $R^M\in C^{1,2}_b(\mathbb{R}_T)$ such that $R^M_x>0$ everywhere,

$$R^M(t,x) := \begin{cases} R(t,x), & x\in[-M,M] \\ 1-r^M\,e^{-x+M}, & x\in[M+1,\infty) \\ r^M\,e^{x+M}, & x\in(-\infty,-M-1] \end{cases} \tag{6.22}$$

and

$$\frac{R^M_t(t,x)^2}{R^M_x(t,x)} \le 2\,\max\left(\frac{R_t(t,M)^2}{R_x(t,M)},\frac{R_t(t,-M)^2}{R_x(t,-M)}\right), \tag{6.23}$$

$$\frac{R^M_{xx}(t,x)^2}{R^M_x(t,x)} \le 2\,\max\left(\frac{R_{xx}(t,M)^2}{R_x(t,M)},\frac{R_{xx}(t,-M)^2}{R_x(t,-M)}\right), \tag{6.24}$$

$$R^M_x(t,x)^3 \le 2\,\max\big(R_x(t,M)^3,R_x(t,-M)^3\big) \tag{6.25}$$

on $[0,T]\times([M,M+1]\cup[-M-1,-M])$. Now, we define

$$h^M=\frac{R^M_t-(A(R^M)R^M_x)_x}{A(R^M)R^M_x} \tag{6.26}$$

pointwise on $\mathbb{R}_T$. It is clear that the paths $\gamma^M(t)=R^M_x(t,x)\,\mathrm{d}x$, $t\in[0,T]$ converge to γ in $C([0,T],M_1(\mathbb{R}))$ and that, for each $M>0$, the function R^M can be chosen

smooth enough to ensure that the restriction of h^M to $[0,T] \times [-M-1, M+1]$ is an element of $C_b(\mathbb{R}_T)$ and Lipschitz. Indeed, note that each function R^M_x is bounded away from 0 on every compact set and the functions R^M_t, R^M_x, R^M_{xx} can be chosen to be Lipschitz on $[0,T] \times [-M-1, M+1]$. Finally,

$$h^M = -\frac{A'(R^M)}{A(R^M)} r^M\, e^{-x+M} + 1 \quad \text{on} \quad [0,T] \times [M+1, \infty), \quad \text{and} \tag{6.27}$$

$$h^M = -\frac{A'(R^M)}{A(R^M)} r^M\, e^{x+M} - 1 \quad \text{on} \quad [0,T] \times (-\infty, -M-1], \tag{6.28}$$

which shows that $h^M \in C_b(\mathbb{R}_T)$ is globally Lipschitz for each $M > 0$.

It remains to show (2.5). To this end, defining h as in the proof of Proposition 2.1, we note that it suffices to prove

$$\limsup_{M\to\infty} \int_{[0,T]\times[M,\infty)} \Big(h^M - \frac{b(R^M)}{A(R^M)}\Big)^2 A(R^M) R^M_x \,\mathrm{d}m = 0, \tag{6.29}$$

$$\limsup_{M\to\infty} \int_{[0,T]\times(-\infty,-M]} \Big(h^M - \frac{b(R^M)}{A(R^M)}\Big)^2 A(R^M) R^M_x \,\mathrm{d}m = 0, \tag{6.30}$$

since

$$\limsup_{M\to\infty} \int_{\mathbb{R}_T\setminus[0,T]\times[-M,M]} \Big(h - \frac{b(R)}{A(R)}\Big)^2 A(R)\, R_x \,\mathrm{d}m = 0 \tag{6.31}$$

due to $J(R) < \infty$. We will only prove the first claim, since the same proof applies to the second claim as well. By the Cauchy-Schwarz inequality and the tightness of the family $R_x(t,x)\,\mathrm{d}x$, $t \in [0,T]$, the former limit superior is bounded above by

$$C \limsup_{M\to\infty} \int_{[0,T]\times[M,\infty)} \frac{(R^M_t)^2}{R^M_x} + \frac{(R^M_{xx})^2}{R^M_x} + (R^M_x)^3 \,\mathrm{d}m \tag{6.32}$$

for some constant $C > 0$ depending only on the supremum of A' and the infimum of A. We now split the respective integrals into an integral over $[0,T]\times[M,M+1]$ and an integral over $[0,T]\times[M+1,\infty)$. The integrals over $[0,T]\times[M,M+1]$ tend to 0 at least for a subsequence of $\{M :\ M \in \mathbb{N}\}$ due to $J(R) < \infty$, (6.23), (6.24) and (6.25). Finally, the integrals over $[0,T] \times [M+1,\infty)$ can be computed to $r^M\, T + (r^M)^3\, \frac{T}{3}$. By decreasing the constants r^M, $M > 0$ if necessary, we can ensure that the latter quantity tends to 0 in the limit $M \to \infty$. This finishes the proof. □

7. Proof of Proposition 2.4

7.1. **Local large deviations upper bound.** The goal of this section is to establish the local large deviations upper bound with the function I. We start with the local large deviations upper bound for paths $\gamma \in \mathcal{A}$. For the definitions of the function I and the space $\mathcal{A}$ we refer the reader to the outline in section 2.

Proposition 7.1. *For each $\gamma \in \mathcal{A}$ the local large deviations upper bound*

$$\lim_{\delta\downarrow 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\rho^N \in B(\gamma,\delta)) \le -I(\gamma) \tag{7.1}$$

holds.

Proof. Step 1. Fix a path $\gamma \in \mathcal{A}$ and a function $g \in \overline{\mathcal{S}}$ as in the definition of the function I. Setting $Z_N^g(t) = (\rho^N(t), g(t,.))$, $t \in [0,T]$ and using Itô's formula, one deduces

$$\mathrm{d}Z_N^g(t) = H^g(t,\rho^N(t))\,\mathrm{d}t + \frac{1}{N}\sum_{i=1}^N \sigma(F_{\rho^N(t)}(X_i(t)))g_x(t,X_i(t))\,\mathrm{d}W_i(t), \tag{7.2}$$

where for each $t \in [0,T]$ and $\alpha \in M_1(\mathbb{R})$ we have set

$$H^g(t,\alpha) = \big(\alpha, g_t(t,.) + b(F_\alpha(.))g_x(t,.) + \frac{1}{2}\sigma(F_\alpha(.))^2 g_{xx}(t,.)\big). \tag{7.3}$$

We note that the quadratic variation process of the martingale part M_N^g of the process Z_N^g can be written as $\frac{1}{N}\int_0^t A^g(\rho^N)(s)\,\mathrm{d}s$, where for each $\xi \in C([0,T],M_1(\mathbb{R}))$ and $t \in [0,T]$ we have put

$$A^g(\xi)(t) = \big(\xi(t), \sigma(F_{\xi(t)}(.))^2(g_x(t,.))^2\big). \tag{7.4}$$

From the Martingale Representation Theorem we can conclude that

$$\mathrm{d}M_N^g(t) = \frac{1}{\sqrt{N}}\sqrt{A^g(\rho^N)(t)}\,\mathrm{d}\beta_N(t) \tag{7.5}$$

for a suitable one-dimensional standard Brownian motion β_N. We now introduce the comparison process Y_N^g given by

$$Y_N^g(t) = \int_0^t H^g(s,\gamma(s))\,\mathrm{d}s + M_N^{g,\gamma}(t), \quad t \in [0,T], \tag{7.6}$$

where

$$M_N^{g,\gamma} = \frac{1}{\sqrt{N}}\int_0^t \sqrt{A^g(\gamma)(s)}\,\mathrm{d}\beta_N(s), \quad t \in [0,T] \tag{7.7}$$

is a martingale.

Step 2. We will show next that the processes $Z_N^g(t)$, $t \in [0,T]$ and $Y_N^g(t)$, $t \in [0,T]$ are close on the event $\rho^N \in B(\gamma,\delta)$ up to a set of negligible probability on the exponential scale. To this end, we fix a $\rho \in B(\gamma,\delta)$, write $\tilde{R}$ for $F_{\rho(.)}(.)$ and R for $F_{\gamma(.)}(.)$, and note

$$\begin{aligned}
&\int_0^T \big|H^g(t,\rho(t)) - H^g(t,\gamma(t))\big|\,\mathrm{d}t \\
\leq\ &\int_0^T \big|\big(\rho(t), b(\tilde{R})g_x + \frac{1}{2}\sigma(\tilde{R})^2 g_{xx} - b(R)g_x - \frac{1}{2}\sigma(R)^2 g_{xx}\big)\big|\,\mathrm{d}t \\
&+ \int_0^T \big|\big(\rho(t), g_t + b(R)g_x + \frac{1}{2}\sigma(R)^2 g_{xx}\big) - \big(\gamma(t), g_t + b(R)g_x + \frac{1}{2}\sigma(R)^2 g_{xx}\big)\big|\,\mathrm{d}t
\end{aligned}$$

and

$$\begin{aligned}\int_0^T \big|A^g(\rho)(t) - A^g(\gamma)(t)\big|\,\mathrm{d}t \;&\leq\; \int_0^T \Big|\big(\rho(t), \sigma(\tilde{R})^2(g_x)^2\big) - \big(\rho(t), \sigma(R)^2(g_x)^2\big)\Big|\,\mathrm{d}t \\ &\quad + \int_0^T \Big|\big(\rho(t), \sigma(R)^2(g_x)^2\big) - \big(\gamma(t), \sigma(R)^2(g_x)^2\big)\Big|\,\mathrm{d}t.\end{aligned}$$

Since, for every $t \in [0,T]$, the measure $\gamma(t)$ has no atoms, we conclude that the function $x \mapsto R(t,x)$ is continuous for every $t \in [0,T]$. Hence, the second summands in both upper bounds tend to zero in the limit $\delta \downarrow 0$ uniformly in $\rho \in B(\gamma,\delta)$. To show that the same is true for the first summands, it suffices to prove

$$\lim_{\delta\downarrow 0} \sup_{\rho\in B(\gamma,\delta)} \int_0^T \int_{\mathbb{R}} |\tilde{R}(t,.) - R(t,.)|\,\mathrm{d}\rho(t)\,\mathrm{d}t = 0, \tag{7.8}$$

since the functions b and σ^2 are Lipschitz (see Assumptions 1, 2). Moreover, for any fixed $\delta > 0$, the definition of the Lévy distance d_L implies that for all $t \in [0,T]$, $\rho \in B(\gamma,\delta)$ and $x \in \mathbb{R}$ one has

$$R(t, x-\delta) - \delta \leq \tilde{R}(t,x) \leq R(t,x+\delta) + \delta \tag{7.9}$$

and, hence, also

$$|\tilde{R}(t,x) - R(t,x)| \leq R(t,x+\delta) - R(t,x-\delta) + 2\delta. \tag{7.10}$$

Thus, the claim (7.8) can be reduced to

$$\lim_{\delta\downarrow 0} \sup_{\rho\in B(\gamma,\delta)} \int_0^T \int_{\mathbb{R}} \big(R(t,x+\delta) - R(t,x-\delta)\big)\,\rho(t)(\mathrm{d}x)\,\mathrm{d}t = 0. \tag{7.11}$$

By applying Fubini's Theorem to the inner integral, we can rewrite the latter double integral as

$$\int_{[0,T]\times\mathbb{R}^2} \mathbf{1}_{(x-\delta,x+\delta]}(y)\,\rho(t)(\mathrm{d}x)\,\gamma(t)(\mathrm{d}y)\,\mathrm{d}t = \int_{[0,T]\times\mathbb{R}^2} \mathbf{1}_{[y-\delta,y+\delta)}(x)\,\rho(t)(\mathrm{d}x)\,\gamma(t)(\mathrm{d}y)\,\mathrm{d}t.$$

Moreover, for any $\rho \in B(\gamma,\delta)$, the definition of the Levy distance d_L yields the inequality

$$\begin{aligned}\int_{\mathbb{R}} \mathbf{1}_{[y-\delta,y+\delta)}(x)\,\rho(t)(\mathrm{d}x) \;&\leq\; \int_{\mathbb{R}} \mathbf{1}_{[y-2\delta,y+2\delta)}(x)\,\gamma(t)(\mathrm{d}x) + \delta \\ &\leq\; R(t,y+3\delta) - R(t,y-3\delta) + \delta.\end{aligned}$$

Hence, (7.8) will follow if we can show

$$\lim_{\delta\downarrow 0} \int_{\mathbb{R}_T} \big(R(t,y+3\delta) - R(t,y-3\delta)\big)\,\gamma(t)(\mathrm{d}y)\,\mathrm{d}t = 0. \tag{7.12}$$

But this follows directly from the Dominated Convergence Theorem, since the functions $R(t,.)$ are continuous for all $t \in [0,T]$ due to $\gamma \in \mathcal{A}$. All in all, we have

shown

$$\lim_{\delta\downarrow 0}\sup_{\rho\in B(\gamma,\delta)}\int_0^T \big|H^g(t,\rho(t))-H^g(t,\gamma(t))\big|\,\mathrm{d}t=0, \tag{7.13}$$

$$\lim_{\delta\downarrow 0}\sup_{\rho\in B(\gamma,\delta)}\int_0^T \big|A^g(\rho)(t)-A^g(\gamma)(t)\big|\,\mathrm{d}t=0. \tag{7.14}$$

From the first identity it follows immediately that for any fixed $\epsilon>0$:

$$\begin{aligned}&\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\rho^N\in B(\gamma,\delta),\ \sup_{t\in[0,T]}|Z_N^g(t)-Y_N^g(t)|>\epsilon\big)\\ \le\ &\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\rho^N\in B(\gamma,\delta),\ \sup_{t\in[0,T]}|M_N^g(t)-M_N^{g;\gamma}(t)|>\epsilon/2\big).\end{aligned}$$

To further estimate the right side, we note that by [31, Theorem 8.5.7] $M_N^g(.)-M_N^{g;\gamma}(.)=B(\tau(.))$ with a suitable standard Brownian motion B and

$$\tau(t)=\int_0^t\langle M_N^g-M_N^{g;\gamma}\rangle(s)\,\mathrm{d}s,\quad t\in[0,T], \tag{7.15}$$

where $\langle M_N^g-M_N^{g;\gamma}\rangle$ is the quadratic variation process of the process $M_N^g-M_N^{g;\gamma}$. Moreover, on the event $\rho^N\in B(\gamma,\delta)$:

$$\tau(t)=\frac{1}{N}\int_0^t\Big(\sqrt{A^g(\rho^N)(s)}-\sqrt{A^g(\gamma)(s)}\Big)^2\,\mathrm{d}s\le\frac{C(\delta)}{N}$$

with a function $C:(0,\infty)\to(0,\infty)$ for which $\lim_{\delta\downarrow 0}C(\delta)=0$. The last inequality follows from the elementary inequality $(\sqrt{x_1}-\sqrt{x_2})^2\le|x_1-x_2|$, $x_1,x_2\in[0,\infty)$ and (7.14). Combining this and Bernstein's inequality for Brownian motion, we deduce

$$\begin{aligned}&\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\rho^N\in B(\gamma,\delta),\ \sup_{t\in[0,T]}|M_N^g(t)-M_N^{g;\gamma}(t)|>\epsilon/2\big)\\ \le\ &\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\sup_{t\in[0,C(\delta)/N]}B(t)>\epsilon/2\big)\\ \le\ &\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\exp\big(-\epsilon^2N/(4C(\delta))\big)=-\infty.\end{aligned}$$

Step 3. From the definition of the metric on the space $C([0,T],M_1(\mathbb{R}))$ and the fact that the Lévy distance and the bounded-Lipschitz metric on $M_1(\mathbb{R})$ generate the same topology, we see that there is a function $r:(0,\infty)\to(0,\infty)$ with $\lim_{\delta\downarrow 0}r(\delta)=0$ and

$$\rho\in B(\gamma,\delta)\Longrightarrow\|(\rho(.),g)-(\gamma(.),g)\|_\infty:=\sup_{t\in[0,T]}|(\rho(t),g)-(\gamma(t),g)|\le r(\delta). \tag{7.16}$$

Hence, using the result of the previous step we can conclude

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}(\rho^N\in B(\gamma,\delta))\\
=\;&\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\rho^N\in B(\gamma,\delta),\sup_{t\in[0,T]}|Z_N^g(t)-Y_N^g(t)|\le\epsilon\big)\\
\le\;&\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\|(\rho^N(.),g)-(\gamma(.),g)\|_\infty\le r(\delta),\sup_{t\in[0,T]}|Z_N^g(t)-Y_N^g(t)|\le\epsilon\big)\\
\le\;&\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\sup_{t\in[0,T]}|(\gamma(t),g)-Y_N^g(t)|\le 2\epsilon\big).
\end{aligned}$$

Step 4. To evaluate the latter upper bound, we will show next that the sequence Y_N^g, $N\in\mathbb{N}$ satisfies a LDP on the space $C([0,T],\mathbb{R})$ with a good rate function I^g. Indeed, define the operator

$$F:\;C([0,\infty),\mathbb{R})\to C([0,T],\mathbb{R}),\quad (Fh)(t)=\int_0^t H^g(s,\gamma(s))\,\mathrm{d}s+h(\overline{A}^g(t)),$$

where $\overline{A}^g(t)=\int_0^t A^g(\gamma)(s)\,\mathrm{d}s$. The time-change formalism for Îto integrals (see e.g. [31, Theorem 8.5.7]) shows that the process Y_N^g can be obtained by applying the operator F to $\frac{1}{\sqrt{N}}\tilde\beta(t)$, $t\ge 0$ with a suitable one-dimensional standard Brownian motion $\tilde\beta$. Since the operator F is continuous with respect to the uniform topology on the space $C([0,\infty),\mathbb{R})$, a combination of Schilder's Theorem ([10, Theorem 5.2.3]) and the Contraction Principle ([10, Theorem 4.2.1]) shows that the sequence Y_N^g, $N\in\mathbb{N}$ satisfies a LDP with the good rate function

$$I^g(f)=\inf_{h:\,F(h)=f}\frac{1}{2}\int_0^{\overline{A}^g(T)}\dot h(t)^2\,\mathrm{d}t. \tag{7.17}$$

In addition, we note that $F(h)=f$ implies

$$\dot h(\overline{A}^g(t))=\frac{\dot f(t)-H^g(t,\gamma(t))}{A^g(\gamma)(t)},\quad t\in[0,T]. \tag{7.18}$$

Thus,

$$I^g(f)=\frac{1}{2}\int_0^T\frac{\big|\dot f(u)-\big(\gamma(u),g_t+b(R)g_x+\frac{1}{2}\sigma(R)^2g_{xx}\big)\big|^2}{(\gamma(u),\sigma(R)^2(g_x)^2)}\,\mathrm{d}u. \tag{7.19}$$

Taking the limit $\epsilon\downarrow 0$ in the final upper bound of step 3 and using the findings of this step, we see that we have shown

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\,\mathbb{P}(\rho^N\in B(\gamma,\delta))\\
&\qquad\le-\frac{1}{2}\int_0^T\frac{\big|\frac{\mathrm{d}}{\mathrm{d}u}(\gamma(u),g)-\big(\gamma(u),g_t+b(R)g_x+\frac{1}{2}\sigma(R)^2g_{xx}\big)\big|^2}{(\gamma(u),\sigma(R)^2(g_x)^2)}\,\mathrm{d}u.
\end{aligned}$$

Taking the infimum over all $g\in\overline{\mathcal{S}}$, we end up with the proposition. □

We now turn to the local large deviations upper bound for paths $\gamma\notin\mathcal{A}$.

Proposition 7.2. *For every path $\gamma \notin \mathcal{A}$ the local large deviations upper bound*

$$\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}(\rho^N\in B(\gamma,\delta))=-\infty \tag{7.20}$$

holds.

In the proof of Proposition 7.2, we distinguish the cases that the path $\gamma \notin \mathcal{A}$ has an atom for some $t \in [0,T]$, does not satisfy the tail estimate in the definition of the space $\mathcal{A}$, or is not absolutely continuous in time in the sense specified in the outline. We will refer to the three cases as Proposition 7.2 (a), Proposition 7.2 (b) and Proposition 7.2 (c), respectively.

Proof of Proposition 7.2 (a). Step 1. If $\gamma(0)$ has an atom, then it is clear that $\mathbb{P}(\rho^N\in B(\gamma,\delta))=0$ for all N large enough. Hence, the statement of the proposition is trivial in this case. For this reason, we will assume from now on that there exists an $s\in(0,T]$, an $a\in\mathbb{R}$ and an $\epsilon>0$ such that $\gamma(s)(\{a\})=\epsilon$. Moreover, we claim that it suffices to consider the case $b\equiv 0$. Indeed, otherwise we introduce the martingale

$$M_{N,b}(t)=\sum_{i=1}^N\int_0^t\frac{b(F_{\rho^N(r)}(X_i(r)))}{\sigma(F_{\rho^N(r)}(X_i(r)))}\,\mathrm{d}W_i(r),\quad t\in[0,T] \tag{7.21}$$

and the corresponding change of measure

$$\frac{\mathrm{d}\mathbb{Q}_{N,b}}{\mathrm{d}\mathbb{P}}=\exp\big(M_{N,b}(T)-\frac{1}{2}\langle M_{N,b}\rangle(T)\big). \tag{7.22}$$

Then, by Hölder's inequality, for every $p,q>1$ with $p^{-1}+q^{-1}=1$:

$$\begin{aligned}\mathbb{P}(\rho^N\in B(\gamma,\delta))&=\mathbb{E}^{\mathbb{Q}_{N,b}}\big[\exp\Big(-M_{N,b}(T)-\frac{1}{2}\langle M_{N,b}\rangle(T)\Big)\mathbf{1}_{\{\rho^N\in B(\gamma,\delta)\}}\big]\\&\le\sup_{\rho\in B(\gamma,\delta)}\exp\Big(N\big(1-\frac{p}{2}\big)\int_{\mathbb{R}_T}b(\tilde R)^2/\sigma(\tilde R)^2\,\mathrm{d}\rho(u)\,\mathrm{d}t\Big)\\&\quad\times\mathbb{E}^{\mathbb{Q}_{N,b}}\Big[\exp\Big(-pM_{N,b}(T)-\frac{p^2}{2}\langle M_{N,b}\rangle(T)\Big)\Big]^{1/p}\mathbb{Q}_{N,b}(\rho^N\in B(\gamma,\delta))^{1/q}\\&\le\exp\Big(N\big(1-\frac{p}{2}\big)\|b/\sigma\|_\infty^2\,T\Big)\,\mathbb{Q}_{N,b}(\rho^N\in B(\gamma,\delta))^{1/q}.\end{aligned}$$

Here, we have used the notation $\tilde R$ for $F_{\rho(.)}(.)$. The latter upper bound reveals that, if we can show

$$\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{Q}_{N,b}(\rho^N\in B(\gamma,\delta))=-\infty, \tag{7.23}$$

this would imply (7.20). For the sake of shorter notation, we will assume $b\equiv 0$ in the following instead of proving (7.23).

Step 2. Next, we fix an $N\in\mathbb{N}$ and a $0<\delta<\epsilon$, and use the definitions of the metric on $C([0,T],M_1(\mathbb{R}))$ and the Lévy distance d_L together with the union bound

to obtain

$$\begin{aligned}\mathbb{P}(\rho^N \in B(\gamma,\delta)) &\leq \mathbb{P}\big(\rho^N(s)([a-\delta,a+\delta]) \geq \epsilon-\delta\big)\\ &= \mathbb{P}\Big(\sum_{i=1}^N \mathbf{1}_{\{|X_i(s)-a|\leq\delta\}} \geq N(\epsilon-\delta)\Big)\\ &\leq \binom{N}{\lceil N(\epsilon-\delta)\rceil} \sup_{u_1,\ldots,u_{\lceil N(\epsilon-\delta)\rceil}} \mathbb{P}\big(|Z_i^{(u_i)}(s)-a|<\delta,\ i=1,2,\ldots,\lceil N(\epsilon-\delta)\rceil\big).\end{aligned}$$

Here, $\lceil N(\epsilon-\delta)\rceil$ is the smallest integer not less than $N(\epsilon-\delta)$, the supremum is taken over all processes $u_1,u_2,\ldots,u_{\lceil N(\epsilon-\delta)\rceil}$ adapted to the filtration generated by the Brownian motions $W_1,W_2,\ldots,W_N$, which take values in the interval $[\inf_{[0,1]}\sigma,\sup_{[0,1]}\sigma]$, and we have set

$$Z_i^{(u_i)}(t) = X_i(0) + \int_0^t u_i(r)\,\mathrm{d}W_i(r), \quad t\in[0,T]. \tag{7.24}$$

Since each Îto integral $\int_0^s u_i(r)\,\mathrm{d}W_i(r)$ is given by the L^2-limit of stochastic integrals of processes adapted to the filtration generated by $W_1,W_2,\ldots,W_N$, which are piecewise constant in time, it suffices to take the supremum above over the space of the latter. For this reason, we now let $u_1,u_2,\ldots,u_{\lceil N(\epsilon-\delta)\rceil}$ be simple processes, which are constant on each of the time intervals $[0,t_1),\ldots,[t_k,s)$ for some $0=t_0<t_1<\cdots<t_k<t_{k+1}=s$ and $k\in\mathbb{N}$. In this case the probability that we are looking to optimize becomes

$$\mathbb{P}\Big(\Big|X_i(0)+\sum_{j=0}^k u_i(t_j)(W_i(t_{j+1})-W_i(t_j))-a\Big|<\delta,\ i=1,2,\ldots,\lceil N(\epsilon-\delta)\rceil\Big). \tag{7.25}$$

Next, we fix an $i\in\{1,2,\ldots,N\}$ and observe that the latter probability conditioned on all appearing random variables except for $u_i(t_k)$ and $W_i(t_{k+1})-W_i(t_k)$ computes to

$$\begin{aligned}&\prod_{m\in\{1,2,\ldots,N\}\backslash\{i\}} \mathbf{1}_{\{|X_m(0)+\sum_{j=0}^k u_m(t_j)(W_m(t_{j+1})-W_m(t_j))-a|<\delta\}}\\ &\times\mathbb{E}\big[\mathbf{1}_{\{|X_i(0)+\sum_{j=0}^k u_i(t_j)(W_i(t_{j+1})-W_i(t_j))-a|<\delta\}}\big|u_i(t_l),W_i(t_{l+1})-W_i(t_l):l=0,\ldots,k-1\big].\end{aligned}$$

The latter expression shows that if $u_i(t_k)$ is not measurable with respect to the σ-algebra generated by $W_i(t)$, $t\in[0,t_k]$, $u_i(t_0),u_i(t_1),\ldots,u_i(t_{k-1})$, then the probability we are looking to optimize can be increased by optimizing the latter conditional expectation over $u_i(t_k)$. Arguing by backward induction we conclude that it suffices to take the latter supremum only over processes $u_1,u_2,\ldots,u_{\lceil N(\epsilon-\delta)\rceil}$ such that the process u_i is adapted to the filtration generated by the Brownian motion W_i for every $i=1,2,\ldots,\lceil N(\epsilon-\delta)\rceil$. Hence, the supremum above can be rewritten as

$$\sup_{u_1}\mathbb{P}(|Z_1^{(u_1)}(s)-a|<\delta)^{\lceil N(\epsilon-\delta)\rceil}, \tag{7.26}$$

where we now take the supremum over all processes u_1 adapted to the filtration generated by the Brownian W_1, which take values in the interval $[\inf_{[0,1]}\sigma, \sup_{[0,1]}\sigma]$. A standard computation involving Stirling's formula yields

$$\lim_{\delta\downarrow 0}\lim_{N\to\infty}\frac{1}{N}\log\binom{N}{\lceil N(\epsilon-\delta)\rceil} = -\epsilon\log\epsilon-(1-\epsilon)\log(1-\epsilon). \tag{7.27}$$

From this and (7.26) we conclude that it suffices to prove

$$\limsup_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\sup_{u_1}\mathbb{P}(|Z_1^{(u_1)}(s)-a|<\delta)^{\lceil N(\epsilon-\delta)\rceil} = -\infty,$$

that is:

$$\limsup_{\delta\downarrow 0}\sup_{u_1}\mathbb{P}(|Z_1^{(u_1)}(s)-a|<\delta) = 0. \tag{7.28}$$

Finally, this latter statement follows immediately from the stronger result in [30, Theorem 1]. □

Proof of Proposition 7.2 (b). Let $\gamma\notin\mathcal{A}$ be such that

$$\int_0^T(\gamma(t),|x|^{1+\eta})\,\mathrm{d}t = \infty. \tag{7.29}$$

Applying the same argument as in step 1 in the proof of Proposition 7.2 (a), we may and will assume $b\equiv 0$. Now, for each $K>1$, let f_K be an infinitely differentiable function taking non-negative values, which satisfies

$$f_K(x)=|x|^{1+\eta}\ \text{ on }\ [-K,-1]\cup[1,K],\quad f_K(x)\le 1\ \text{ on }\ [-1,1], \tag{7.30}$$

$$|f_K'|^2\le 4f_K,\ \ \|f_K\|_\infty\le 2K^{1+\eta},\ \ \|f_K'\|_\infty\le 2K^{1-\eta},\ \ \|f_K''\|_\infty\le 2. \tag{7.31}$$

Such functions can be constructed by smoothing the continuous functions

$$|x|^{1+\eta}\,\mathbf{1}_{[-K,K]}+K^{1+\eta}\,\mathbf{1}_{(-\infty,-K]\cup[K,\infty)} \tag{7.32}$$

around the points $-K$, 0 and K.

Next, we fix an $R>0$, introduce the stopping times

$$\tau_{N,K,R}:=\inf\big\{t\in[0,T]:\ (\rho^N(t),f_K)\ge R\big\} \tag{7.33}$$

and recall $\sup_{N\in\mathbb{N}}(\rho^N(0),|x|^{1+\eta})<\infty$ (see Assumption 1). From the latter inequality it follows that there is a constant $C_0>0$ such that

$$\sup_{K>1}\sup_{N\in\mathbb{N}}(\rho^N(0),f_K)\le C_0. \tag{7.34}$$

Hence, applying the Itô's formula, we may conclude
(7.35)

$$\mathbb{P}(\tau_{N,K,R}\le T)\le\mathbb{P}\Big(\frac{1}{N}\sum_{i=1}^N\int_0^{\tau_{N,K,R}\wedge T}f_K'(X_i(t))\sigma(F_{\rho^N(t)}(X_i(t)))\,\mathrm{d}W_i(t)\ge R-\Delta_K\Big),$$

where we have set

$$\Delta_K=C_0+\frac{1}{2}\|f_K''\|_\infty\|\sigma\|_\infty^2 T. \tag{7.36}$$

Writing the random variable appearing in the latter probability as

$$\beta\Big(\frac{1}{N^2}\sum_{i=1}^N\int_0^{\tau_{N,K,R}\wedge T} f_K'(X_i(t))^2\sigma(F_{\rho^N(t)}(X_i(t)))^2\,\mathrm{d}t\Big) \tag{7.37}$$

for a suitable standard Brownian motion β and using $|f_K'|^2\le 4f_K$, we deduce

$$\begin{aligned}\mathbb{P}(\tau_{N,K,R}\le T) &\le \mathbb{P}\big(\sup_{0\le t\le 4N^{-1}RT\|\sigma\|_\infty^2}\beta(t)\ge R-\Delta_K\big)\\ &\le \exp\Big(-N\frac{(R-\Delta_K)^2}{4RT\|\sigma\|_\infty^2}\Big),\quad R>\sup_{K>1}\Delta_K.\end{aligned}$$

Now, we recall the fact that the Lévy distance and the bounded Lipschitz metric

$$d_{BL}(\alpha_1,\alpha_2):=\sup_{f:\,\|f\|_\infty+\mathrm{L}ip(f)\le 1}\Big|\int_{\mathbb{R}} f\,\mathrm{d}\alpha_1-\int_{\mathbb{R}} f\,\mathrm{d}\alpha_2\Big| \tag{7.38}$$

generate the same topology on $M_1(\mathbb{R})$, where we wrote $\mathrm{L}ip(f)$ for the Lipschitz constant of f (with the convention $\mathrm{L}ip(f)=\infty$ if f is not globally Lipschitz). Therefore, there is a function r taking positive values, for which $\lim_{\delta\downarrow 0} r(\delta)=0$ and such that for any $\alpha_1,\alpha_2\in M_1(\mathbb{R})$ and any bounded Lipschitz function $f:\mathbb{R}\to\mathbb{R}$:

$$\Big|\int_{\mathbb{R}} f\,\mathrm{d}\alpha_1-\int_{\mathbb{R}} f\,\mathrm{d}\alpha_2\Big|\le(\|f\|_\infty+\mathrm{L}ip(f))\,r(d_L(\alpha_1,\alpha_2)). \tag{7.39}$$

Thus, in view of the previous estimate, we have

$$\begin{aligned}&\mathbb{P}\big(\rho^N\in B(\gamma,\delta)\big)\\ \le\;& \mathbb{P}\Big(T^{-1}\int_0^T(\rho^N(t),f_K)\,\mathrm{d}t\ge T^{-1}\int_0^T(\gamma(t),f_K)\,\mathrm{d}t-(2K^{1+\eta}+2K^{1-\eta})r(\delta)\Big)\\ \le\;& \exp\Big(-N\frac{\big(T^{-1}\int_0^T(\gamma(t),f_K)\,\mathrm{d}t-(2K^{1+\eta}+2K^{1-\eta})r(\delta)-\Delta_K\big)^2}{4\big((T^{-1}\int_0^T(\gamma(t),f_K)\,\mathrm{d}t-(2K^{1+\eta}+2K^{1-\eta})r(\delta)\big)T\,\|\sigma\|_\infty^2}\Big)\end{aligned}$$

for all K large enough and $\delta>0$ small enough. It follows that

$$\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\rho^N\in B(\gamma,\delta)\big)\le-\frac{\big(T^{-1}\int_0^T(\gamma(t),f_K)\,\mathrm{d}t-\Delta_K\big)^2}{4\int_0^T(\gamma(t),f_K)\,\mathrm{d}t\,\|\sigma\|_\infty^2}. \tag{7.40}$$

By taking the limit $K\to\infty$ and using $\sup_{K>1}\Delta_K<\infty$ and (7.29), we end up with the desired estimate. □

Proof of Proposition 7.2 (c). We proceed as in the proof of Proposition 7.1 to conclude that (7.20) must hold, if there is a function $g\in\overline{\mathcal{S}}$ such that the map $t\mapsto(\gamma(t),g)$ is not absolutely continuous on $[0,T]$, as a consequence of the fact that the rate of not absolutely continuous trajectories is infinite in Schilder's Theorem. □

7.2. **Exponential tightness.** The following proposition establishes the exponential tightness of the sequence ρ^N, $N \in \mathbb{N}$.

Proposition 7.3. *The sequence ρ^N, $N \in \mathbb{N}$ is exponentially tight on $C([0,T], M_1(\mathbb{R}))$ in the sense that for every $M > 0$ there exists a compact set $K_M \subset C([0,T], M_1(\mathbb{R}))$ for which*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\rho^N \notin K_M) \leq -M. \tag{7.41}$$

Proof. Step 1. We prove the proposition by checking the criterion for exponential tightness in [9, Lemma A.2]. To this end, we need to show that for all rational $t \in [0,T]$ the sequence $\rho^N(t)$, $N \in \mathbb{N}$ is exponentially tight on $(M_1(\mathbb{R}), d_L)$, and that for any fixed $\epsilon > 0$ one has

$$\lim_{\delta\downarrow 0} \sup_{N\in\mathbb{N}} \frac{1}{N} \log \mathbb{P}\Big(\sup_{0\leq s,t\leq T, |t-s|\leq\delta} d_L(\rho^N(s), \rho^N(t)) > \epsilon \Big) = -\infty. \tag{7.42}$$

The first assertion is shown in step 2 and the second one in step 3. Here, d_L stands for the symmetrized Lévy distance defined in the introduction.

Step 2. Consider the function $\phi : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$. By Prokhorov's Theorem the set

$$\{\alpha \in M_1(\mathbb{R}) :\ (\alpha, \phi) \leq C\} \tag{7.43}$$

is pre-compact in $(M_1(\mathbb{R}), d_L)$ for any $C > 0$. Hence, to prove the first assertion, it is enough to show that, for each rational $t \in [0,T]$ and any $M > 0$, one can find a $C > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}((\rho^N(t), \phi) > C) \leq -M. \tag{7.44}$$

To this end, we first apply Markov's inequality to obtain for any $C > 0$:

$$\begin{aligned}
&\mathbb{P}((\rho^N(t), \phi) > C) = \mathbb{P}\Big(\frac{1}{N} \sum_{i=1}^N |X_i(t)| > C \Big) \\
\leq\ & \mathbb{P}\Big(\frac{1}{N} \sum_{i=1}^N \Big| \int_0^t \sigma(F_{\rho^N(s)}(X_i(s)))\, \mathrm{d}W_i(s) \Big| > C - \|b\|_\infty T - (\rho^N(0), \phi) \Big) \\
\leq\ & e^{-(C - \|b\|_\infty T - (\rho^N(0),\phi))\, N} \sup_{u_1,\ldots,u_N} \mathbb{E}\Big[\prod_{i=1}^N e^{|Z^{(u_i)}(t)|} \Big].
\end{aligned}$$

Here, the supremum is taken over all processes $u_1, \ldots, u_N$ adapted to the filtration generated by the Brownian motions $W_1, \ldots, W_N$, which take values in the interval $[\inf_{[0,1]} \sigma, \sup_{[0,1]} \sigma]$, and we have set

$$Z^{(u_i)}(t) = \int_0^t u_i(s)\, \mathrm{d}W_i(s), \quad t \in [0,T]. \tag{7.45}$$

Arguing as in step 2 in the proof of Proposition 7.2 (a), we conclude that it suffices to take the supremum over processes $u_1,\ldots,u_N$ such that, for each $i=1,2,\ldots,N$, the process u_i is adapted to the filtration generated by the Brownian motion W_i. Thus, the upper bound on $\mathbb{P}((\rho^N(t),\phi)>C)$ simplifies to

$$e^{-(C-\|b\|_\infty T-(\rho^N(0),\phi))N}\sup_{u_1}\mathbb{E}\Big[e^{|Z^{(u_1)}(t)|}\Big]^N, \tag{7.46}$$

where now the supremum is taken over all processes u_1, which are adapted to the filtration generated by W_1 and take values in the interval $[\inf_{[0,1]}\sigma,\sup_{[0,1]}\sigma]$. We conclude

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}((\rho^N(t),\phi)>C)\le-(C-\|b\|_\infty T-(\rho_0,\phi))+\log\sup_{u_1}\mathbb{E}\Big[e^{|Z^{(u_1)}(t)|}\Big]$$

due to Assumptions 1, 2. By viewing $Z^{(u_1)}$ as the result of a time change applied to a standard Brownian motion β, we deduce

$$\sup_{u_1}\mathbb{E}\Big[e^{|Z^{(u_1)}(t)}\Big]\le\mathbb{E}\Big[e^{\sup_{0\le r\le\|\sigma\|_\infty^2 t}|\beta(r)|}\Big]. \tag{7.47}$$

The latter two estimates yield (7.44).

Step 3. It remains to show (7.42). From the proof of [12, Theorem 11.3.3] we deduce that for all $0\le s,t\le T$:

$$\begin{aligned} d_L(\rho^N(s),\rho^N(t)) &\le 2\Big(\frac{1}{N}\sup_{f:\,\|f\|_\infty+\mathrm{L}ip(f)\le1}\Big|\sum_{i=1}^N(f(X_i(s))-f(X_i(t)))\Big|\Big)^{\frac12}\\ &\le 2\Big(\frac{1}{N}\sum_{i=1}^N|X_i(t)-X_i(s)|\Big)^{\frac12}. \end{aligned}$$

Combining this with

$$|X_i(t)-X_i(s)|\le\Big|\int_s^t\sigma(F_{\rho^N(r)}(X_i(r)))\,\mathrm{d}W_i(r)\Big|+\|b\|_\infty\,|t-s|, \tag{7.48}$$

we see that it suffices to prove

$$\lim_{\delta\downarrow0}\sup_{N\in\mathbb{N}}\frac{1}{N}\log\mathbb{P}\Big(\sup_{0\le s,t\le T,|t-s|\le\delta}\frac{1}{N}\sum_{i=1}^N\Big|\int_s^t\sigma(F_{\rho^N(r)}(X_i(r)))\,\mathrm{d}W_i(r)\Big|>\epsilon\Big)=-\infty$$

for any fixed $\epsilon>0$. In fact, we will show

$$\lim_{\delta\downarrow0}\sup_{N\in\mathbb{N}}\frac{1}{N}\log\mathbb{P}\Big(\frac{1}{N}\sum_{i=1}^N\sup_{0\le s,t\le T,|t-s|\le\delta}\Big|\int_s^t\sigma(F_{\rho^N(r)}(X_i(r)))\,\mathrm{d}W_i(r)\Big|>\epsilon\Big)=-\infty \tag{7.49}$$

for all $\epsilon > 0$. Using Markov's inequality and following the arguments of step 2, we can bound the left side of (7.49) from above by

$$\begin{aligned} &\lim_{\delta\downarrow 0}\sup_{N\in\mathbb{N}}\frac{1}{N}\log\Big(e^{-\lambda N\epsilon}\,\mathbb{E}\Big[\exp\Big(\lambda\sup_{u_1}\sup_{0\le s,t\le T,|t-s|\le\delta}\big|Z^{(u_1)}(t)-Z^{(u_1)}(s)\big|\Big)\Big]^N\Big)\\ =\;& -\epsilon\lambda+\lim_{\delta\downarrow 0}\log\mathbb{E}\Big[\exp\Big(\lambda\sup_{u_1}\sup_{0\le s,t\le T,|t-s|\le\delta}\big|Z^{(u_1)}(t)-Z^{(u_1)}(s)\big|\Big)\Big]\end{aligned}$$

with arbitrary $\lambda > 0$ and where the supremum is taken over all processes u_1 taking values in the interval $[\inf_{[0,1]}\sigma, \sup_{[0,1]}\sigma]$ and adapted to the filtration generated by W_1. Viewing $Z^{(u_1)}$ again as the result of a time change applied to a standard Brownian motion β, we can estimate the latter upper bound further by

$$-\epsilon\lambda+\lim_{\delta\downarrow 0}\log\mathbb{E}\Big[\exp\Big(\lambda\sup_{0\le s,t\le T\|\sigma\|_\infty^2,|t-s|\le\delta\|\sigma\|_\infty^2}|\beta(t)-\beta(s)|\Big)\Big].$$

Now, combining the bound

$$\sup_{0\le s,t\le T\|\sigma\|_\infty^2,|t-s|\le\delta\|\sigma\|_\infty^2}|\beta(t)-\beta(s)|\le 2\sup_{0\le t\le T\|\sigma\|_\infty^2}|\beta(t)| \tag{7.50}$$

with the Dominated Convergence Theorem, we see that the second summand in the latter upper bound is equal to zero. In other words, we have shown

$$\lim_{\delta\downarrow 0}\sup_{N\in\mathbb{N}}\frac{1}{N}\log\mathbb{P}\Big(\frac{1}{N}\sum_{i=1}^N\sup_{0\le s,t\le T,|t-s|\le\delta}\Big|\int_s^t\sigma(F_{\rho^N(r)}(X_i(r)))\,\mathrm{d}W_i(r)\Big|>\epsilon\Big)\le-\epsilon\lambda \tag{7.51}$$

for all $\lambda > 0$. Thus, we can finish the proof by taking the limit $\lambda\to\infty$. □

8. Proof of Proposition 2.5

In this section we show the local large deviations lower bound for paths $\gamma\in\mathcal{G}$ and, thus, complete the proof of the LDP.

Proof of Proposition 2.5. Step 1. We may assume that $J(\gamma)<\infty$. Fix a $\gamma\in\mathcal{G}$ and an $N\in\mathbb{N}$. We write R for $F_{\gamma(\cdot)}(\cdot)$ as before and recall the existence of a Lipschitz function $h\in C_b(\mathbb{R}_T)$ of (1.11) given explicitly as

$$h=\frac{R_t-(A(R)R_x)_x}{A(R)R_x} \tag{8.1}$$

due to positivity of R_x. Next, we introduce the processes

$$M_N(t)=\sum_{i=1}^N\int_0^T\Big(\frac{1}{2}\,h(t,X_i(t))\sigma(F_{\rho^N(t)}(X_i(t)))+\frac{b(F_{\rho^N(t)}(X_i(t)))}{\sigma(F_{\rho^N(t)}(X_i(t)))}\Big)\,\mathrm{d}W_i(t), \tag{8.2}$$

$$A_N(t)=\sum_{i=1}^N\int_0^T\Big(\frac{1}{2}\,h(t,X_i(t))\sigma(F_{\rho^N(t)}(X_i(t)))+\frac{b(F_{\rho^N(t)}(X_i(t)))}{\sigma(F_{\rho^N(t)}(X_i(t)))}\Big)^2\,\mathrm{d}t \tag{8.3}$$

on $[0,T]$. By Girsanov's Theorem there is a unique probability measure $\mathbb{Q}$ given by

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}=\exp(M_N(T)-A_N(T)/2), \tag{8.4}$$

and under $\mathbb{Q}$

(8.5)
$$\mathrm{d}X_i(t) = -\frac{1}{2}h(t,X_i(t))\sigma(F_{\rho^N(t)}(X_i(t)))^2\,\mathrm{d}t + \sigma(F_{\rho^N(t)}(X_i(t)))\,\mathrm{d}W_i^{\mathbb{Q}}(t),\ i=1,\ldots,N,$$

where $W_1^{\mathbb{Q}},\ldots,W_N^{\mathbb{Q}}$ are independent standard Brownian motions. Moreover, we observe the identity

$$\frac{1}{2}A_N(T) = \frac{N}{2}\int_{\mathbb{R}_T}\Big(\frac{1}{2}\,h(t,x)\sigma(F_{\rho^N(t)}(x)) + \frac{b(F_{\rho^N(t)}(x))}{\sigma(F_{\rho^N(t)}(x))}\Big)^2\,\rho^N(t)(\mathrm{d}x)\,\mathrm{d}t. \tag{8.6}$$

Now, letting $\tilde{R} := F_{\rho(.)}(.)$, $\tilde{h} := \frac{1}{2}\,h\sigma(R) + \frac{b(R)}{\sigma(R)}$ and $\tilde{h}_\rho := \frac{1}{2}\,h\sigma(\tilde{R}) + \frac{b(\tilde{R})}{\sigma(\tilde{R})}$, where ρ is a generic element of the ball $B(\gamma,\delta)$, and applying the triangle inequality, we deduce

$$\begin{aligned}
&\sup_{\rho\in B(\gamma,\delta)}\Big|\int_{\mathbb{R}_T}\tilde{h}_\rho^2\,\rho(t)(\mathrm{d}x)\,\mathrm{d}t - \int_{\mathbb{R}_T}\tilde{h}^2\,\gamma(t)(\mathrm{d}x)\,\mathrm{d}t\Big| \\
\leq\ &\sup_{\rho\in B(\gamma,\delta)}\Big|\int_{\mathbb{R}_T}\tilde{h}_\rho^2\,\rho(t)(\mathrm{d}x)\,\mathrm{d}t - \int_{\mathbb{R}_T}\tilde{h}^2\,\rho(t)(\mathrm{d}x)\,\mathrm{d}t\Big| \\
&+\sup_{\rho\in B(\gamma,\delta)}\Big|\int_{\mathbb{R}_T}\tilde{h}^2\,\rho(t)(\mathrm{d}x)\,\mathrm{d}t - \int_{\mathbb{R}_T}\tilde{h}^2\,\gamma(t)(\mathrm{d}x)\,\mathrm{d}t\Big|.
\end{aligned}$$

We claim that the first term in the latter upper bound converges to 0 in the limit $\delta\downarrow 0$. Indeed, opening the square and using the boundedness of h, b and σ, the fact that the values of σ are bounded away from 0 and the Lipschitz property of b and σ^2, one can reduce the statement to the identity (7.8). Moreover, the convergence of the second summand to 0 in the limit $\delta\downarrow 0$ can be shown by moving the supremum and the absolute value inside the time integral and applying the Dominated Convergence Theorem. All in all, we conclude that for any fixed $\epsilon>0$:

$$\Big|\frac{1}{2}A_N(T) - N\cdot J(\gamma)\Big| = \Big|\frac{1}{2}A_N(T) - \frac{N}{2}\int_{\mathbb{R}_T}\tilde{h}^2\,\gamma(t)(\mathrm{d}x)\,\mathrm{d}t\Big| \leq N\epsilon \tag{8.7}$$

on the event $\rho^N\in B(\gamma,\delta)$ for all $\delta>0$ small enough.

Step 2. Hölder's inequality and the result of step 1 show that, for any fixed $\epsilon>0$, all $\delta>0$ small enough and all $p,q>1$ with $p^{-1}+q^{-1}=1$, we have the estimates

$$\begin{aligned}
&\mathbb{P}(\rho^N\in B(\gamma,\delta)) = \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(M_N^{\mathbb{Q}}(T) - \frac{1}{2}A_N(T)\Big)\mathbf{1}_{\{\rho^N\in B(\gamma,\delta)\}}\Big] \\
&\geq \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(-\frac{q}{p}M_N^{\mathbb{Q}}(T) + \frac{q}{2p}A_N(T)\Big)\Big]^{-p/q}\mathbb{Q}\big(\rho^N\in B(\gamma,\delta)\big)^p \\
&= \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(-\frac{q}{p}M_N^{\mathbb{Q}}(T) - \frac{q^2}{2p^2}A_N(T)\Big)e^{(\frac{q}{2p}+\frac{q^2}{2p^2})A_N(T)}\Big]^{-p/q}\mathbb{Q}\big(\rho^N\in B(\gamma,\delta)\big)^p \\
&\geq e^{-(1+\frac{q}{p})N(J(\gamma)+\epsilon)}\,\mathbb{Q}\big(\rho^N\in B(\gamma,\delta)\big)^p,
\end{aligned}$$

where

$$M_N^{\mathbb{Q}}(t) = -\sum_{i=1}^N\int_0^T\Big(\frac{1}{2}\,h(t,X_i(t))\sigma(F_{\rho^N(t)}(X_i(t))) + \frac{b(F_{\rho^N(t)}(X_i(t)))}{\sigma(F_{\rho^N(t)}(X_i(t)))}\Big)\,\mathrm{d}W_i^{\mathbb{Q}}(t). \tag{8.8}$$

The latter lower bound reveals that it suffices to show

$$\lim_{N\to\infty} \mathbb{Q}\big(\rho^N \in B(\gamma,\delta)\big) = 1, \tag{8.9}$$

since then the statement of the proposition follows from

$$\lim_{\epsilon\downarrow 0}\lim_{p\uparrow\infty,q\downarrow 1}\lim_{\delta\downarrow 0}\lim_{N\to\infty}\frac{1}{N}\log\Big[e^{-(1+\frac{q}{p})N(J(\gamma)+\epsilon)}\,\mathbb{Q}\big(\rho^N\in B(\gamma,\delta)\big)^p\Big] = -J(\gamma). \tag{8.10}$$

Step 3. To prove (8.9), for any fixed $N\in\mathbb{N}$, we let $Q_{N,h}$ be the law of ρ^N under $\mathbb{Q}$. Then, arguing as in steps 1 and 2 in the proof of [37, Theorem 1.1], one deduces the tightness of the sequence $Q_{N,h}$, $N\in\mathbb{N}$. Indeed, the proof given there relies on the completely general compactness criterion for subsets of $C([0,T],M_1(\mathbb{R}))$ given in [17, Lemma 1.3] and can be carried out mutatis mutandis by using the boundedness of the drift and diffusion coefficients in equation (8.5) in place of the boundedness of the corresponding coefficients in the dynamics treated in [37, Theorem 1.1]. Moreover, the computations involving the initial conditions can be omitted here, since $\rho^N(0)\to\rho_0$, $N\to\infty$ (see Assumption 1). We therefore conclude that, in order to prove (8.9), it suffices to show that every limit point Q_h of the sequence $Q_{N,h}$, $N\in\mathbb{N}$ is given by the Dirac measure δ_γ.

As will become apparent from the proof, we may assume without loss of generality that Q_h is the limit of the whole sequence $Q_{N,h}$, $N\in\mathbb{N}$. Moreover, in view of the Skorokhod Representation Theorem in the form of [13, Theorem 3.5.1], we may also assume that the random variables ρ^N, $N\in\mathbb{N}$ are defined on the same probability space and converge to a limiting random variable $\tilde\rho$ of law Q_h almost surely in $C([0,T],M_1(\mathbb{R}))$. Fixing a $g\in\overline{\mathcal{S}}$ we see by means of Îto's formula that

$$(\rho^N(t),g(t,\cdot)) - (\rho^N(0),g(0,\cdot)) = \int_0^t (\rho^N(s),(L_{h,\rho^N}g)(s,\cdot))\,\mathrm{d}s + M_N^g(t),\quad t\in[0,T] \tag{8.11}$$

where we have set

$$L_{h,\rho} = \frac{\partial}{\partial t} - \frac{1}{2}h\,\sigma(\tilde R)^2\,\frac{\partial}{\partial x} + \frac{1}{2}\sigma(\tilde R)^2\frac{\partial^2}{\partial x^2}, \tag{8.12}$$

$$M_N^g(t) = \frac{1}{N}\sum_{i=1}^N\int_0^t g_x(s,X_i(s))\sigma(F_{\rho^N(s)}(X_i(s)))\,\mathrm{d}B_i(s),\quad t\in[0,T] \tag{8.13}$$

with ρ being a generic element of $C([0,T],M_1(\mathbb{R}))$, $\tilde R = F_{\rho(.)}(.)$ as before and appropriate independent standard Brownian motions $B_1,B_2,\dots,B_N$.

It is clear that the left side in equation (8.11) converges to the random variable $(\tilde\rho(t),g(t,.)) - (\tilde\rho(0),g(t,.))$ for all $t\in[0,T]$ with probability one. Moreover, in Lemma 8.1 below we will show that the measures $\tilde\rho(t)$, $t\in[0,T]$ have no atoms with probability 1. Thus, the function $\tilde R = F_{\tilde\rho(.)}(.)$ is continuous almost surely. Hence, arguing as in step 2 in the proof of Proposition 7.1 and using the continuity and

boundedness of all functions involved, one deduces that

$$\lim_{N\to\infty}\sup_{0\le t\le T}\Big|\int_0^t (\rho^N(s),(L_{h,\rho^N}g)(s,.))\,\mathrm{d}s-\int_0^t(\tilde\rho(s),(L_{h,\tilde\rho}g)(s,.))\,\mathrm{d}s\Big|=0 \tag{8.14}$$

with probability 1. Furthermore, the boundedness of the functions g_x and σ implies that the quadratic variation $\langle M_N^g\rangle(T)$ of the martingale M_N^g at time T converges to 0. Thus, the L^2-version of Doob's maximal inequality for non-negative submartingales implies that $\sup_{0\le t\le T}|M_N^g(t)|$ converges to zero in the L^2-sense as $N\to\infty$. Combining the latter three observations, we conclude that the equation

$$(\tilde\rho(t),g(t,.))-(\tilde\rho(0),g)=\int_0^t(\tilde\rho(s),(L_{h,\tilde\rho}g)(s,.))\,\mathrm{d}s \tag{8.15}$$

must hold for all $t\in[0,T]$ with probability 1. Therefore, the latter equation holds for all $t\in[0,T]$ and for all g belonging to a countable dense subset of $\overline{\mathcal{S}}$ with probability 1. Thus, integrating by parts with respect to the spatial variable and recalling $\frac{1}{2}\sigma(\tilde R)=A(\tilde R)$, we can conclude that $\tilde R=F_{\tilde\rho(.)}(.)$ is a weak solution of the partial differential equation

$$\tilde R_t=(A(\tilde R)\tilde R_x)_x+h\,A(\tilde R)\tilde R_x. \tag{8.16}$$

However, in the upcoming Lemma 8.2 we show that the weak solution of the latter equation is unique. Hence, $\tilde R=R$ and $\tilde\rho=\gamma$ must hold with probability 1. □

To complete the proof of Proposition 2.5, we need to show the following two lemmas.

Lemma 8.1. *Let the sequence $Q_{N,h}$, $N\in\mathbb{N}$ be defined as in step 3 in the proof of Proposition* 2.5*. Then, for every limit point Q_h of this sequence and every random path $\tilde\rho$ of law Q_h, the measures $\tilde\rho(t)$, $t\in[0,T]$ do not have atoms with probability* 1.

Proof. As will become clear from the proof, we may assume that Q_h is the limit of the whole sequence $Q_{N,h}$, $N\in\mathbb{N}$. Moreover, in view of the Skorokhod Representation Theorem in the form of [13, Theorem 3.5.1], we may assume as before that the random variables ρ^N, $N\in\mathbb{N}$ are defined on the same probability space and converge to the random path $\tilde\rho$ almost surely in $C([0,T],M_1(\mathbb{R}))$. Arguing as in steps 2 and 3 in the proof of Proposition 7.2 (a) below, we see that, for all $\epsilon>0$, there exists a $\delta>0$ such that for all $s\in[0,T]$ and $a\in\mathbb{R}$:

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{Q}\big(\rho^N(s)([a-\delta,a+\delta])>\epsilon-\delta\big)\le-1. \tag{8.17}$$

Furthermore, we can follow the arguments in step 2 in the proof of Proposition 7.3 to deduce that, for all $\epsilon>0$, there exists a $\delta>0$ and a compact interval $K\subset\mathbb{R}$ such that, for all $s\in[0,T]$:

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{Q}\big(\rho^N(s)(\mathbb{R}\backslash K)>\epsilon-\delta\big)\le-1. \tag{8.18}$$

Next, fix an $\epsilon > 0$, and let $\delta > 0$ and $K \subset \mathbb{R}$ be such that (8.17) and (8.18) hold. Now, fix points $a_1, a_2, \ldots, a_l$ in K such that every point in K is at most at distance $\frac{\delta}{2}$ from the set $\{a_1, a_2, \ldots, a_l\}$. Combining (8.17) for $a \in \{a_1, \ldots, a_l\}$, (8.18) and the Borel-Cantelli Lemma, we conclude that, for any fixed $s \in [0, T]$:

$$\sup_{a\in\mathbb{R}} \rho^N(s)([a-\delta/2, a+\delta/2]) \le \epsilon - \delta \tag{8.19}$$

for all N large enough with probability 1. Clearly, the same is true simultaneously for all s in a given finite set $\{s_1, \ldots, s_m\} \subset [0, T]$.

Next, arguing as in step 3 in the proof of Proposition 7.3, we see that there is a $\kappa > 0$ such that

$$\sup_{N\in\mathbb{N}} \frac{1}{N} \log \mathbb{Q}\Big(\sup_{0\le s\le t\le T, |s-t|\le\kappa} d_L(\rho^N(s), \rho^N(t)) > \delta/4 \Big) \le -1. \tag{8.20}$$

Using the Borel-Cantelli Lemma again, we see that

$$d_L(\rho^N(s), \rho^N(t)) \le \delta/4 \tag{8.21}$$

for all $0 \le s \le t \le T$ with $|s-t| \le \kappa$ and all N large enough, with probability 1. Finally, we choose the number $m \in \mathbb{N}$ and the set $\{s_1, \ldots, s_m\} \subset [0, T]$ such that

$$\sup_{t\in[0,T]} \min_{j=1,2,\ldots,m} |t - s_i| < \kappa. \tag{8.22}$$

Picking for any $t \in [0, T]$ an element $s(t) \in \{s_1, \ldots, s_m\}$ with $|s(t) - t| < \kappa$, we see that

$$\rho^N(t)([a-\delta/4, a+\delta/4]) \le \rho^N(s(t))([a-\delta/2, a+\delta/2]) + \delta/4 \le \epsilon - \delta + \delta/4 \le \epsilon$$

for all $t \in [0, T]$, all $a \in \mathbb{R}$ and all $N \in \mathbb{N}$ large enough, with probability 1. Combining this with the Portmanteau Theorem, we infer that, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\begin{aligned} &\sup_{t\in[0,T]} \sup_{a\in\mathbb{R}} \tilde{\rho}(t)((a-\delta/4, a+\delta/4)) \le \sup_{t\in[0,T]} \sup_{a\in\mathbb{R}} \liminf_{N\to\infty} \rho^N(t)((a-\delta/4, a+\delta/4)) \\ \le\ & \limsup_{N\to\infty} \sup_{t\in[0,T]} \sup_{a\in\mathbb{R}} \rho^N(t)((a-\delta/4, a+\delta/4)) \le \epsilon. \end{aligned}$$

This finishes the proof of the lemma. □

Lemma 8.2. *For every Lipschitz function $h \in C_b(\mathbb{R}_T)$, the weak solution of the Cauchy problem*

$$R_t = (A(R)R_x)_x + h\, A(R) R_x, \tag{8.23}$$
$$R(0, .) = F_{\rho_0} \tag{8.24}$$

in the space $\mathcal{R} = \{R \in C_b(\mathbb{R}_T) :\ R_x(., x)\,\mathrm{d}x \in C([0,T], M_1(\mathbb{R}))\}$ is unique.

Proof. We adapt the proof of [11, Theorem 4.2 (1)] to the present situation. To this end, we assume first that $h_x \leq 0$ Lebesgue almost everywhere and will treat the general case below. We note that [11, Proposition 2.1] applies without alteration to the equation (8.23) and asserts the existence of a bounded weak solution u to the Cauchy problem (8.23), (8.24), which is the pointwise limit of classical solutions u_j, $j \in \mathbb{N}$ of the equation (8.23) (indeed, the proof of [11, Proposition 2.1] relies merely on [28, chapter V, Theorem 8.1] and on the maximum principle for quasilinear parabolic equations, which both apply to the equation (8.23)).

Next, we introduce the concept of subsolutions and supersolutions for the equation (8.23). Recalling the definition of the function Σ in (1.4), we call a $w \in \mathcal{R}$ a subsolution of the equation (8.23) if the inequality

$$\int_{\mathbb{R}} w(t,x)g(t,x)\,\mathrm{d}x - \int_{\mathbb{R}} w(0,x)g(0,x)\,\mathrm{d}x \leq \int_{\mathbb{R}_t} w\,g_t + \Sigma(w)\big(g_{xx} - h\,g_x - h_x\,g\big)\,\mathrm{d}m$$

is satisfied for all $t \in [0,T]$ and all $g \in \overline{\mathcal{S}}$ taking non-negative values. Here, we wrote $\mathbb{R}_t$ for $[0,t] \times \mathbb{R}$. We call w a supersolution if the inequality holds in the reversed direction. To prove the lemma, it suffices to show that there exists an absolute constant $c > 0$ such that

$$\int_{\mathbb{R}} (u(t,x) - w(t,x))_+ \omega(x)\,\mathrm{d}x \leq c \int_{\mathbb{R}} (u(0,x) - w(0,x))_+ e^{-|x|}\,\mathrm{d}x \tag{8.25}$$

holds for any supersolution w of (8.23), any $t \in [0,T]$ and any $\omega \in C_c^\infty(\mathbb{R})$ taking values in $[0,1]$ (the same inequality with $u-w$ replaced by $w-u$ for a subsolution w is then shown in an analogous manner, so that the uniqueness of solutions to (8.23) in $\mathcal{R}$ follows).

Proceeding as in the proof of [11, Theorem 4.2] we see that (8.25) will follow if we can prove the estimates (i)-(v) of [11, Lemma 4.1] for the unique classical solutions of the first boundary value problems

$$\begin{aligned} \zeta_t + A_{n,j}\zeta_{xx} + B_{n,j}\zeta_x + C_{n,j}\zeta = 0 \quad \text{on} \;\; (0,t) \times (-r,r),& \\ \zeta(t,x) = \omega(x)\chi(x),\;\; x \in [-r,r],& \\ \zeta(s,-r) = \zeta(s,r) = 0,\;\; s \in [0,t],& \end{aligned} \tag{8.26}$$

where $r \in (0,\infty)$ and $t \in [0,T]$ are fixed, for each $j \in \mathbb{N}$ the functions $A_{n,j}$, $B_{n,j}$, $C_{n,j}$, $n \in \mathbb{N}$ are smooth approximations of

$$A_j = \int_0^1 A(\tau u_j + (1-\tau)w)\,\mathrm{d}\tau, \tag{8.27}$$

$$B_j = -h \int_0^1 A(\tau u_j + (1-\tau)w)\,\mathrm{d}\tau, \tag{8.28}$$

$$C_j = -h_x \int_0^1 A(\tau u_j + (1-\tau)w)\,\mathrm{d}\tau \tag{8.29}$$

such that $A_{n,j} \downarrow A_j$, $B_{n,j} \downarrow B_j$ uniformly on $[0,t] \times [-r,r]$, $0 \le C_{n,j} \to C_j$ Lebesgue almost everywhere on $[0,t] \times [-r,r]$, and χ is an arbitrary infinitely differentiable function supported in $[-r,r]$ and taking values in $[0,1]$. The existence and uniqueness of the classical solution to (8.26) is well-known (see e.g. [28, chapter IV, Theorem 5.2]).

The estimate (i) of [11, Lemma 4.1] holds for ζ, since $C_{n,j} \ge 0$ and, hence, the maximum principle applies to the equation (8.26) (see e.g. [16, section 2.1, Theorem 1] and note that our time direction is reversed compared to the setting there).

The estimates (ii) and (iii) of [11, Lemma 4.1] for ζ can be deduced by exactly the same arguments as in [11], where one only needs to notice that the maximum principle applies to the equation (8.26) and that in our case the functions $A_{n,j}$, $B_{n,j}$ and $C_{n,j}$ can be chosen to be bounded uniformly in j and n and such that $A_{n,j}$ is bounded away from zero uniformly in j and n (this follows from the fact that the corresponding properties hold for the functions A_j, B_j and C_j uniformly in j).

The validity of estimate (iv) of [11, Lemma 4.1] for ζ is a direct consequence of [28, chapter III, Theorem 11.1]. Finally, estimate (v) of [11, Lemma 4.1] can be shown to hold for ζ by following the lines of the proof in [11] and estimating the additional term

$$\int_0^t \int_{-r}^r (C_{n,j}\zeta)\zeta_{xx} \,\mathrm{d}x\,\mathrm{d}s, \tag{8.30}$$

which will appear on the right side of their equation (4.10), by applying the Cauchy-Schwarz inequality.

Arguing as on page 394 of [11] we deduce from the latter five estimates

$$\int_{\mathbb{R}} (u_j(t,x) - w(t,x))\omega(x)\chi(x)\,\mathrm{d}x \le c \int_{\mathbb{R}} (u_j(0,x) - w(0,x))_+ e^{-|x|}\,\mathrm{d}x, \tag{8.31}$$

where $c > 0$ is the constant in the estimate (ii). The value of c can be chosen independently of j, since in our case the functions A_j, B_j and C_j are bounded uniformly in j (see also the corresponding remark at the end of the proof of [11, Lemma 4.1 (ii)]). Our refined estimate (8.31) replacing the estimate (4.12) in [11] is needed, since the implication (4.12)⇒(4.13) there is not justified.

Choosing functions χ which approximate the function $\mathbf{1}_{\{u_j > w\}}$ pointwise, we conclude from (8.31):

$$\int_{\mathbb{R}} (u_j(t,x) - w(t,x))_+ \omega(x)\,\mathrm{d}x \le c \int_{\mathbb{R}} (u_j(0,x) - w(0,x))_+ e^{-|x|}\,\mathrm{d}x. \tag{8.32}$$

Finally, we take the limit $j \to \infty$ and apply the Dominated Convergence Theorem to obtain (8.25) from (8.32). This finishes the proof in the case that $h_x \le 0$ holds. In the general case, it is not hard to see that we can apply the same proof to the partial differential equation satisfied by the function $U = e^{-vt}\,R$ for any fixed number $v > \|h_x\|_\infty \, \|A\|_\infty$. □

Acknowledgement. We thank N. Krylov for suggesting the method applied in the proof of Lemma 4.4 (which is one of the key ingredients in the proof of Proposition 2.2 (A)), and for describing to one of us some of the details needed for the implementation of the method.

References

[1] Aronson, D. G. (1968). Non-negative solutions of linear parabolic equations. *Ann. Scuola Superiore Pisa, Classe di Scienze 3e série* **22** 607-694.

[2] Aubin J.-P., Ekeland I. (1984). *Applied nonlinear analysis*. John Wiley & Sons, New York.

[3] Bass R., Pardoux E. (1987). Uniqueness for diffusions with piecewise constant coefficients. *Probab. Theory Related Fields* **76** 557-572.

[4] Bossy M., Talay D. (1996). Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation. *Annals of Applied Probability* **6** 818-861.

[5] Bossy M., Talay D. (1997). A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Mathematics of Computation* **66** 157-192.

[6] Chatterjee S., Pal S. (2010). A phase transition behavior for Brownian motions interacting through their ranks. *Probab. Theory Relat. Fields* **147** 123-159.

[7] Chatterjee S., Pal S. (2008). A combinatorial analysis of interacting diffusions. To appear in *J. Theor. Probab.*.

[8] Dawson D. A., Gärtner J. (1987). Large deviations from the McKean-Vlasov Limit for Weakly Interacting Diffusions. *Stochastics* **20** 247-308.

[9] Dembo A., Zajic T. (1995). Large deviations: From empirical mean and measure to partial sums processes. *Stoch. Proc. Appl.* **57** 191-224.

[10] Dembo A., Zeitouni O. (1998). *Large deviation techniques and applications*. 2nd ed. Springer, New York.

[11] Diaz J. I., Kersner R. (1987). On a Nonlinear Degenerate Parabolic Equation in Infiltration or Evaporation through a Porous Medium. *J. Differential Equations* **69** 368-403.

[12] Dudley R. M. (2002). *Real analysis and probability*. 2nd ed. Cambridge University Press, New York.

[13] Dudley R. M. (1999). *Uniform central limit theorems*. Cambridge University Press.

[14] Fernholz E. R. (2002). *Stochastic Portfolio Theory*. Springer, New York.

[15] Fernholz R., Karatzas, I. (2009). Stochastic Portfolio Theory: an Overview. In: Bensoussan, A., Zhang, Q. (eds.) *Handbook of Numerical Analysis: Volume XV: Mathematical Modeling and Numerical Methods in Finance*, 89-167. North Holland, Oxford.

[16] Friedman A. (1964). *Partial differential equations of parabolic type*. Prentice-Hall, Inc. Englewood Cliffs, N. J.

[17] Gärtner J. (1988). On the McKean-Vlasov Limit for Interacting Diffusions. *Math. Nachr.* **137** 197-248.

[18] Gilding B. H. (1989). Improved theory for a nonlinear degenerate parabolic equation. Ann. Sc. Norm. Super. Pisa Cl. Sci. **16** 165-224.

[19] Ichiba T., Karatzas, I. (2009). On collisions of Brownian particles. *Ann. Appl. Probab.* **20** 951-977.

[20] Ichiba T., Karatzas I., Shkolnikov M. (2011). Strong solutions of stochastic equations with rank-based coefficients. To appear in *Probab. Theory Related Fields*. Available at *http://arxiv.org/abs/1109.3823*.

[21] Ichiba T., Papathanakos V., Banner A., Karatzas I., Fernholz R. (2010). Hybrid Atlas Models. *Ann. Appl. Probab.* **21** 609-644.

[22] Jourdain B. (2000). Diffusion processes associated with nonlinear evolution equations for signed measures. *Methodology and Computing in Applied Probability* **2:1** 69-91.

[23] Karatzas I., Shreve S. (1998). *Brownian motion and stochastic calculus*. 2nd ed. Springer, New York.

[24] Krylov N. V. (1982). *Controlled diffusion processes*. Springer, Berlin.

[25] Krylov N. V. (1976). Sequences of convex functions and bounds on the maximum of the solution to parabolic equations. *Siberian Math. J.* **17** 290-303.

[26] Krylov N. V. (2007). Parabolic and Elliptic Equations with VMO Coefficients. *Comm. Partial Differential Equations* **32** 453-475.

[27] Krylov N. V. (2012). Some L_p-estimates for elliptic and parabolic operators with measurable coefficients. *Discrete and Continuous Dynamical Systems Series B* **17** 2073-2090.

[28] Ladyzhenskaja O. A., Solonnikov V. A. and Ural'ceva N. N. (1988). *Linear and quasilinear equations of parabolic type*. Translations of Mathematical Monographs **23**. American Mathematical Society.

[29] Lieberman G. (1996). *Second order parabolic differential equations*. World Scientific. River Edge, N.J.

[30] McNamara J. M. (1985). A regularity condition on the transition probability measure of a diffusion process. *Stochastics* **15** 161-182.

[31] Oksendal B. K. (2003). *Stochastic differential equations*. 6th ed. Springer, Berlin-Heidelberg.

[32] Pal S., Pitman J. (2008). One-dimensional Brownian particle systems with rank-dependent drifts. *Ann. Appl. Probab.* **18** 2179-2207.

[33] Pal S., Shkolnikov M. (2011). Concentration of measure for systems of Brownian particles interacting through their ranks. Preprint available at *arXiv:1011.2443v1*.

[34] Portenko N. I. (1974). On the solutions of stochastic differential equations with integrable drift coefficient. *Probability Theory and Its Applications* **XIX**, 577-582.

[35] Rockafeller R. T. (1970). *Convex analysis*. Princeton University Press.

[36] Sato K. (2002). *Levy processes and infinitely divisible distributions*. Engl. ed. Cambridge University Press, Cambridge.

[37] Shkolnikov M. (2010). Large systems of diffusions interacting through their ranks. To appear in *Stoch. Proc. Appl.*. Available at *arXiv:1008.4611*.

[38] Srivastava S. M. (1998). *A course on Borel sets*. Graduate Texts in Mathematics, 180. Springer, New York.

[39] Stroock D. W., Varadhan S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, New York.

[40] Stroock D. W., Varadhan S. R. S. (1967). Diffusion processes with continuous coefficients II. *Comm. Pure Appl. Math.* **xxii**, 479-530.

[41] Stroock D. W., Varadhan S. R. S. (1979). *Multidimensional diffusion processes*. Springer, Berlin.

[42] Vázquez, J. L. (2007). *The porous medium equation. Mathematical theory.* Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford.

Departments of Statistics and Mathematics, Stanford University, Stanford, CA 94305
E-mail address: adembo@stat.stanford.edu

Department of Statistics, University of California, Berkeley, CA 94720-3860
E-mail address: mshkolni@gmail.com

Courant Institute, New York University, New York, NY 10012-1185
E-mail address: varadhan@cims.nyu.edu

School of Mathematics, University of Minnesota, Minneapolis, MN 55455, and Faculty of Mathematics, Weizmann Institute of Science, Rehovot, Israel 76100
E-mail address: zeitouni@umn.edu